UNIVERSITA' DEGLI STUDI DI VERONA

Dipartimento di Biotecnologie

Dottorato di ricerca in
Biotecnologie Molecolari, Industriali ed Ambientali

Ciclo XXII

NMR study of chicken Liver Bile Acid Binding Protein: interaction and dynamics.

S.S.D. CHIM/06

Coordinatore: Prof. Roberto Bassi

Tutors:

Prof. Henriette Molinari (University of Verona)

Dr. Laura Ragona (ISMAC-CNR, Milano)

Dottoranda: Dott.ssa Clelia Cogliati

Contents

Abbreviations

ASBT	Apical Sodium-Dependent Bile Acid Transporter
BA	Bile Acid
BSEP	Bile Acids Export Pump
BMRB	Biological Magnetic Resonance Bank
CA	Cholic Acid
CDA	Chenodeoxycholic acid
cL-BABP	chicken Liver Bile Acid Binding Protein
cL-BABP/S-S	chicken Liver Bile Acid Binding Protein with disulphide bridge
CMC	Critical micellar Concentrations
CPMG	Carr-Purcell- Meiboom-Gill
CRABP	Cellular Retinol Binding Protein
CSP	Chemical Shift Perturbation
CSD	Cross-Strand Disulfide
DIPSI	Decoupling In the Presence of Scalar Interactions
DOSY	Diffusion Ordered Spectroscopy
DTT	Dithiotreithol
FABP	Fatty Acid Binding Protein
FATP5	Fatty Acid Transport Protein 5
FPLC	Fast Protein Purification Chromatography
FXR	Farnesoid X receptor
GCA	Glycocholic Acid
GCDA	Glycochenodeoxycholic acid
Hsp33	Heat Shock Hsp33 C
HSQC	Heteronuclear Single Quantum Coherence
I-BABP	Ileal Bile Acid Binding Protein
iLBPs	intracellular Lipid Binding Proteins
I-LBP	Ileal Lipid-Binding Protein
Lb-FABP	Liver basic Fatty Acid Binding Proteins
L-BABP	Liver Bile Acid Binding Protein
mEH	microsomal Epoxide Hydrolase
MRP2	Multidrug Resistance Protein 2

MRP3 and MRP4	Multidrug Resistance Proteins 3 and 4
NMR	Nuclear Magnetic Resonance
NTCP	Na^+-taurocholate Cotransporting Polypeptide
OATP	Organic Anion Transporting Polypeptide
OST	Organic Solute Transporter
OxyR,	DNA-binding transcriptional dual regulator
PMSF	Phenylmethanesulfonylfluoride
RD	Relaxation Dispersion
SCOP	database containing disulphide bridge classification in protein families
SDS-page	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
S-S	Disulphide bridge
TALOS	Torsion Angle Likelihood Obtained from Shift and sequence similarity
TCA	Taurocholic Acid

Abstract

The aim of this thesis is to understand the role played by a naturally occurring disulphide bridge on the bile acid (BA) binding and functional properties of cytosolic Liver Bile Acid Binding Protein (L-BABP). Bile acids circulate between liver and intestine through a mechanism defined as "enterohepatic circulation", which is a tightly regulated process, particularly by BAs themselves. Indeed BAs are able to influence the expression of numerous genes involved in their synthesis and transport by binding to the primary intracellular nuclear bile acid receptor, farnesoid X receptor (FXR). Understanding the mechanism regulating the interactions of intracellular carriers with bile acid is a key step to provide a model for the transfer of BAs from cytoplasm to the nucleus and can be used to inspire design of therapeutic agents in the treatment of metabolic disorders, such as obesity, type 2 diabetes, hyperlipidaemia and atherosclerosis.

To achieve a detailed molecular and dynamical description of the binding mechanism driving to the formation of the ternary complex of L-BABPs with two BA molecules, spectroscopic methods together with kinetic and thermodynamic analysis have been applied and implemented. In particular structural, dynamical and interaction properties of two forms of chicken L-BABP (cL-BABP), differing by the presence/absence of a naturally occurring disulphide bridge, have been investigated through nuclear magnetic resonance (NMR) approaches. The study of protein-ligand interactions by NMR was performed analysing complexes where, alternatively, either the protein or the ligand were isotopically labelled. ^{15}N enriched glycocholic (GCA) and glycochenodeoxycholic acid (GCDA), two of the most important members of bile salts pool, were employed for protein titrations and their resonances followed through the acquisition and analysis of several NMR experiments (HSQC, DOSY). The obtained results shed light on binding stoichiometry and ligand exchange phenomena but were not sufficient to derive detailed information on affinity, cooperativity and binding mechanism. Thus NMR lineshape analysis as a function of ligand concentration was chosen as an appropriate tool to investigate the complex interaction mechanism within the cL-BABP/BA system. In this line, new NMR approaches have been recently described which allow a reliable and sensitive investigation of ligand binding events occurring on microsecond to millisecond (μs-ms) time scales using lineshape and relaxation dispersion

experiments[1]. Particularly, the combination of these NMR methods can be useful in the study of complex multi-step mechanisms, allowing the correlation between protein dynamics and function[2].

^{15}N relaxation studies, performed on the apo-protein, revealed the presence of slow motions occurring on the microseconds-milliseconds timescale. The central question to be addressed is here whether these motions are essential for ligand uptake, how they can eventually lead to conformations competent for binding and how they are influenced by the presence of the disulfide bridge.

The analysis of titration experiments of ^{15}N labelled protein with unlabelled GCDA through lineshape analysis and relaxation dispersion allowed to define a multi-step binding mechanism for bile salt binding to liver BABPs and to provide an estimate of the kinetics involved.

Summary of the work

The main question addressed in this thesis work is related to the identification of the role played by a disulphide bridge on the structural, dynamic and binding properties of cL-BABP/S-S*.

To this aim the following steps had to be accomplished:

Set up of the protocol for high yield expression and purification of cL-BABP/S-S (unlabelled, singly (^{15}N) and doubly (^{13}C, ^{15}N) labelled samples):

- Assignment of the backbone resonances of cL-BABP/S-S in its apo form, in homotypical complexes with chenodeoxycholic, glycochenodeoxycholic and cholic acids (CDA, GCDA, GCA) and in a heterotypical complex with both GCDA and GCA;
- Study of the backbone dynamics of cL-BABP/S-S in its apo form, in homotypical complexes with CDA, GCDA, GCA and in a heterotypical complex with both GCDA and GCA, through measurement of ^{15}N T_1, T_2 and ^{1}H-^{15}N Nuclear Overhauser Effects (NOE);
- Study of the interaction mode of cL-BABP/S-S with bile acid glycoderivatives through NMR titration experiments, exploiting either protein or ligand observation;
- Lineshape analysis simulations applied to a series of titration experiments for the identification of the kinetic parameters defying the binding mechanism;
- Analysis of the relaxation dispersion data to investigate a “conformational selection” model of binding.

*As far as the nomenclature of the mentioned proteins is concerned, the protein with a disulphide bridge was initially referred to as cL-BABP T91C. This nomenclature has been recently corrected to L-BABP/S-S and L-BABP to underline the existence of both forms in chicken liver.

1. Introduction

To better understand biological functions of bile acid transporters and the rules regulating their interactions within the cell, it is here briefly described the physiological context within which they are inserted.

1.1 Enterohepatic circulation

Liver hepatocytes metabolize cholesterol to cholic acid and chenodeoxycholic acid. These lipid-soluble bile acids are conjugated mainly to glycine or taurine molecules to form water soluble primary conjugated bile acids. After conjugation bile acids are secreted into the canalicular space between hepatocytes via two ATP-binding cassette transporters: bile acids export pump (BSEP), that carries monovalent bile acids[3] and multidrug resistance protein 2 (MRP2), that transports divalent sulfated or glucuronidated molecules[4]. Thus bile acids are secreted against steep concentration gradients in the gallbladder, together with cholesterol and phospholipids. The gallbladder stores about 50 ml of bile, which is released when food containing fat enters the digestive tract. Through the bile duct, bile acids reach the intestinal lumen, where they carry out their role in emulsifying dietary lipids. Some of the secreted bile acids undergo bacterial deconjugation in the distal small intestine and in the colon, becoming then membrane permeable. The transformation of bile acids by bacterial enzymes promotes passive absorption in the colon of those molecules escaping the active uptake that takes place in the ileum. If both mechanisms operate normally, only 1%-3% of the amount of bile acids secreted by the liver is ultimately excreted in faeces (deconjugated and otherwise transformed). While unconjugated bile acids may passively diffuse across the small intestinal and colonic epithelia, conjugated bile acids are actively absorbed in the distal ileum via the apical sodium-dependent bile acid transporter (ASBT), which imports bile acids coupled to Na^+ absorption in a 1:2 stoichiometry[5]. Bile acids species are not equally transported by ASBT: conjugated (more hydrophilic) forms are preferred to unconjugated forms. This is physiologically consistent with the fact that deconjugation normally takes place in the colon. The affinity of ASBT is also higher for

dihydroxy (CDCA and DCA) than for trihydroxy (CA, taurocholic acid-TCA, glycocholic acid-GCA) bile acids[6]. Bile acids are believed to be transferred directly from ASBT to the ileal bile acid-binding protein (I-BABP) through the formation of a 2:1 stoichiometric complex[7]. I-BABP is supposed to facilitate transport of bile acids within the enterocyte to the basolateral membrane. The basolateral efflux of bile acids from enterocytes into the portal vein is mediated by the organic solute transporter α and β (OSTα/OSTβ)[8]. Once in the portal vein, bile acids associate with albumin and reach the space of Disse through the large fenestrae of the liver sinusoids. After their dissociation from albumin, conjugated bile acids uptake into hepatocytes is mediated by Na^+-taurocholate cotransporting polypeptide (NTCP), a basolateral sodium-dependent cotransporter that shares homology with ASBT[9], [10]. Uptake may also be mediated by one or more of the multiple sodium independent anion transporters belonging to the organic anion transporting polypeptide (OATP) family and by the microsomal epoxide hydrolase (mEH)[11], [12]. The mechanisms by which unconjugated bile acids enter the hepatocyte have not been clarified yet. Four possible mechanisms are likely: i) transport mediated by the fatty acid transport protein 5 (FATP5)[13]; ii) transport via NTCP, since at least one unconjugated bile acid (UDCA) has been shown to be a substrate for this carrier[14]; iii) transport via the multiple (sodium-independent) OATP carriers; iv) passive flipflop across the basolateral membrane[15]. Conjugated bile acids are also transported from the hepatocytes back into sinusoidal blood via the multidrug resistance proteins 3 and 4 (MRP3 and MRP4), that are cotransporters of bile acids and glutathione. There is thus uptake, regurgitation, and further uptake of bile acids as bile acids move down the sinusoid[16].

A schematic representation of bile acids enterohepatic circulation is shown in Figure 1.1.

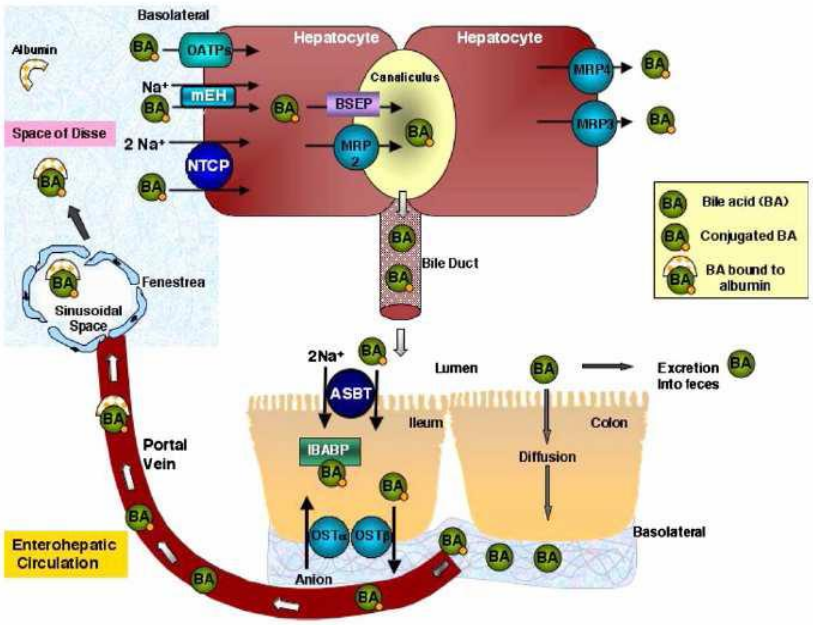
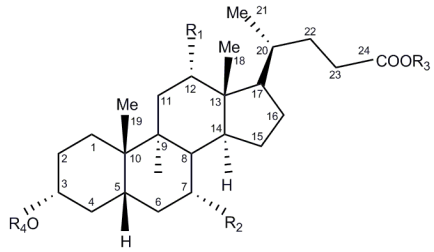


Figure 1.1 Enterohepatic circulation of bile acids [12].

1.2 Bile Acids

Bile acids are amphipathic steroidal compounds biosynthesized from the enzymatic catabolism of cholesterol in the liver, are stored in gallbladder and are the main components of bile.

All primary bile acids appear to have three features in common: (i) they are the major end-products of cholesterol metabolism, (ii) they are secreted into the bile largely in a conjugated form, and (iii) these conjugates are membrane-impermeable, water-soluble, amphiphilic molecules having a powerful ability to transform lamellar arrays of lipids into mixed micelles[17], [18], [19], [20], [21]. The bile acid skeleton, illustrated in Figure 1.2, consists of a steroid body (nucleus) and a tail (side chain). Both parts of the bile acid molecule have a large number of possible polar substituents and conjugating groups may be present on the nucleus (e.g. sulphate, glucuronate, phosphate) or on the side chain (glycine, taurine or other amino acids, and even sugars).



	R_1	R_2
Cholic Acid	OH	OH
Chenodeoxycholic Acid	H	OH
Deoxicholic acid	OH	H

Figure 1.2 Schematic representation of a bile acid molecule. The most common substituents for positions C-12, C-7, are reported in Table (R_1 and R_2). For position C-24 and C-3 the substituents are: R_3 = Taurine (-NH-CH_2-CH_2-SO_3H), glycine (-NH-CH_2-CO_2H), or hydrogen (H); R_4 = Sulphuric acid (-HSO_4), glucuronic acid (-$C_6H_9O_7$) or hydrogen (H), respectively.

Both bile acid chemistry and nomenclature are complex because of the great variety of chemical structures in naturally occurring compounds. The great number of different bile acids can be explained by the evolution of several biochemical pathways converting cholesterol, a poorly soluble membrane lipid, into conjugated bile acids, which are highly water-soluble molecules with amphipathic and membranolytic properties.
Bile acids from different species chemically differ in three respects: (i) side-chain structure, (ii) stereochemistry of steroid rings fusion, and (iii) the distribution of number, position and stereochemistry of hydroxyl groups in the steroid nucleus.

1.2.1 Bile acid biosynthesis

Cholesterol degradation to bile acids in the liver can be initiated by either cholesterol 7α-hydroxylase (CYP7A1) in the classic (neutral) pathway, or by mitochondrial sterol 27-hydroxylase (CYP27A1) in the alternative (acidic or sterol 27-hydroxylase) pathway. In the classic pathway, modification of the sterol nucleus, including saturation of the double bond, epimerization of the 3β-hydroxyl group, and hydroxylation at the 7α and 12α- positions, precedes oxidative cleavage of the side chain. In the alternative pathway, side chain oxidation precedes steroid ring modification. While the classic pathway of bile acids biosynthesis is limited to hepatocytes, the alternative pathway exists in all tissues.

In Figure 1.3 a schematic version of these complex metabolic pathways is reported.

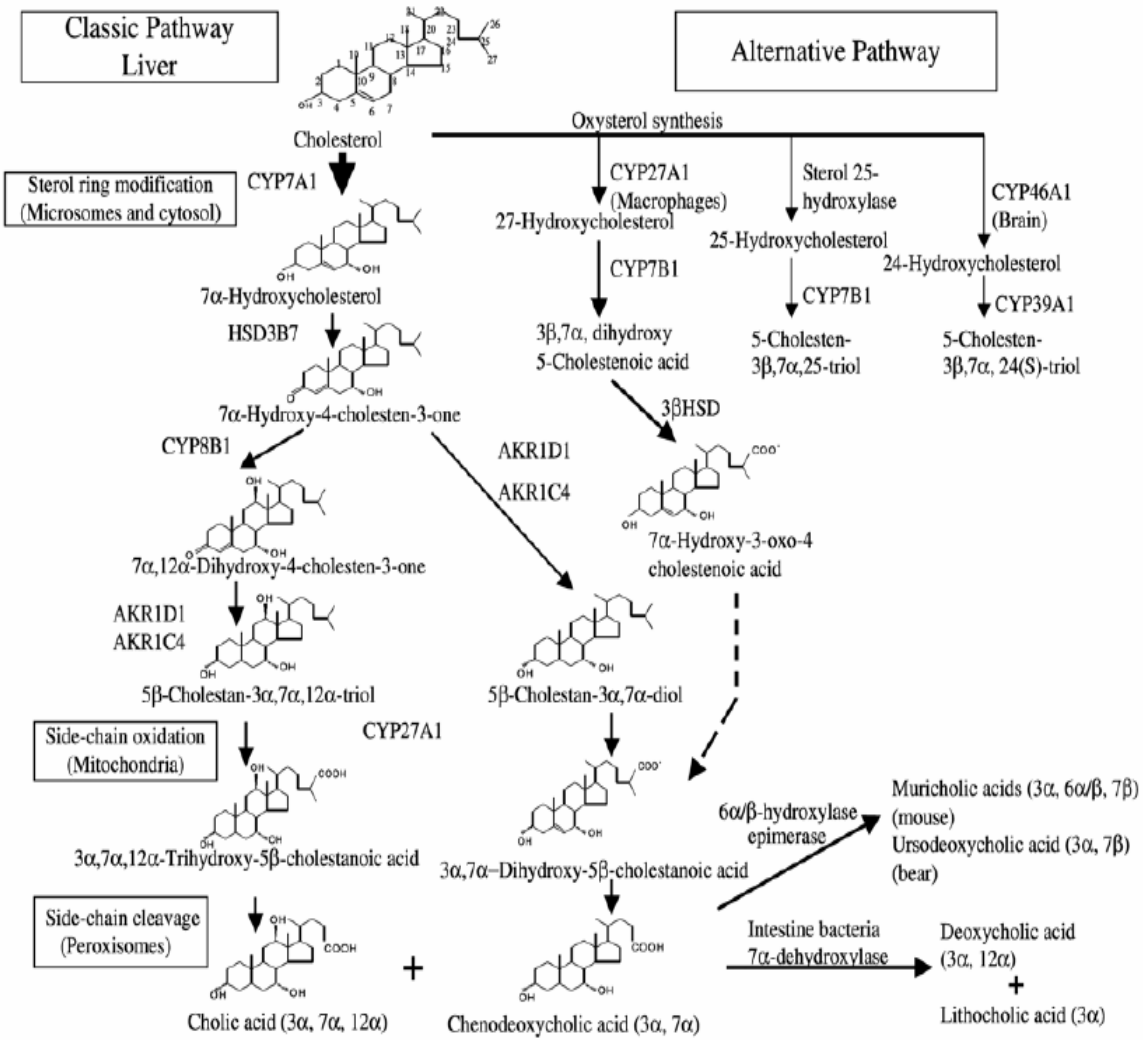


Figure 1.3 Bile acids biosynthetic pathways. Only major regulatory steps and enzymes are shown[22].

This biosynthesis is characterised by a number of complex steps involving modification in both the steroid nucleus and the side chain. In the first step, believed to be the rate limiting step, cholesterol is oxidized to 7α-hydroxycholesterol by cholesterol 7α-hydroxylase. 7α-hydroxycholesterol is then converted to the key intermediate cholest-7α-hydroxy-Δ^4-3-one through the action of an isomerase and a reductase. This unsaturated oxo derivative is the branching point for cholic (CA) and chenodeoxycholic acid (CDA) biosynthesis. Hydroxylation at C-12 leads to the formation of the cholic acid backbone. This oxo derivative is then stereoselectively reduced to afford the 5β-bile acid skeleton. The second key metabolic step, 27-hydroxylation followed by the formation of a C27 carboxylic acid, is believed to occur in the mitochondrion mediated by a P-450 hydroxylase. Then the oxidative cleavage of the side chain (C27 acid to C24

acid) mediated by peroxisomal enzymes affords the mature C24 bile acids. Bile acids are bio-transformed into glycine and taurine conjugates by the amidation reaction catalysed by an acyltransferase.

The relative contribution of the classic and alternative pathways to overall bile acids synthesis is not clear. The classic pathway may be the main process under physiological conditions, whereas the alternative pathway may become the major biosynthetic pathway in patients with liver diseases[22].

1.2.2 Micelle formation

In aqueous environments, bile salts aggregate to form micelles[23], [24], [25]. Bile acids perform almost all physiological functions in the form of mixed-micelles. Bile-salt micelles can solubilise cholesterol, lecithin, monoglycerides, etc., which are intrinsically water-insoluble[26]. The aqueous solubility of cholesterol (~1 nM) can increase more than a million fold in the presence of bile-salt micelles[27], [28]. The cholesterol solubilisation ability is far better with dihydroxy bile-salts (i.e.: chenodeoxycholate salt) than with trihydroxy bile-salts (i.e.: cholate salt).

Bile acids are facially amphipathic, they contain both hydrophobic (lipid soluble) and hydrophilic (polar) faces. The rigid steroid backbone has polar hydroxyl groups on the concave α-face and methyl groups on the convex β-face, see Figure 1.4.

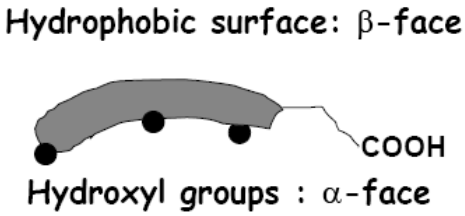


Figure 1.4 Cartoon representation of the rigid steroic backbone of a bile acid introduced by Small[25].

This unique facial amphiphilicity of this class of molecules, enable them to aggregate in aqueous media in a manner different from conventional detergents. Aggregation of bile salts in aqueous solution is largely driven by the hydrophobic association of apolar β-faces of steroid backbones, while further aggregation occurs through hydrogen bonding interactions (Figure 1.5). Critical micellar concentrations (CMCs) of dihydroxy bile salts are typically below 5 mM, whereas trihydroxy bile salts have higher CMC ranging

from 10 to 15 mM. The higher CMC of trihydroxy bile salts is attributed to their higher solubility in water. The aggregation number of these bile-salt micelles ranges from 2 to 10 for globular primary aggregates, which increases as a function of the bile-salt concentration (and added salt concentration) to form larger secondary aggregates. Emulsification of fat through mixed micelle formation is one of the significant properties of bile salts, enabling fat digestion and absorption through the intestinal wall[26].

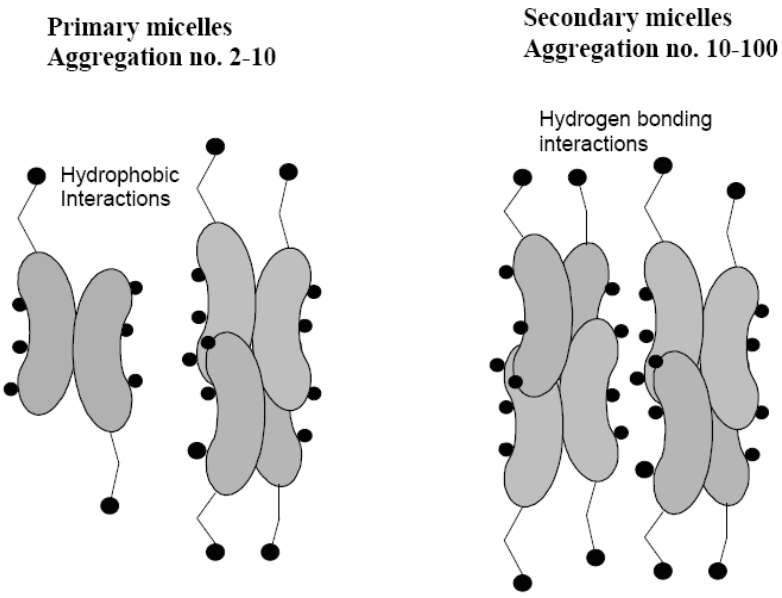


Figure 1.5 Cartoon representations of bile salt micelles (primary and secondary aggregation model) as introduced by Small[25].

1.3 Bile Acid Binding Proteins

Liver basic Fatty Acid Binding Proteins (Lb-FABPs), more accurately described as Liver Bile Acid Binding Proteins (L-BABP)[29] are members of the intracellular Lipid Binding Proteins (iLBPs) family, involved in the transport and "enterohepatic circulation" of bile acids as well as in the regulation of cholesterol metabolism.

FABPs have been classified and described on the basis of the organ that they were initially isolated from, but several instances are known in which more than one FABP type have been shown to be produced by a single tissue. A higher similarity of L-FABPs from non mammalian species with mammalian Ileal Lipid-Binding Proteins (I-LBPs) rather than with mammalian liver FABPs is reported[29]. In agreement with this observation, bile acid binding and transport are emerging as the specific function of the

liver non mammalian subfamily, hence called liver bile acid binding protein[30]. At variance, the paralogous proteins expressed in the same tissue but in mammals play a role in fatty acids binding and transport[31]. A multiple alignment of liver BABP, isolated from chicken liver, with ileal BABPs belonging to different species is reported in Figure 1.6

Figure 1.6 ClustalW multiple alignment of cL-BABP with ileal BABP sequences. For the ileal proteins only the name of the species is reported. cL-BABP residues involved in ligand binding are highlighted in bold[32].

Intracellular LBPs (iLBPs) are a family of phylogenetically related low molecular weight (14-15 kDa) proteins. Although various functions have been proposed for iLBPs, it is widely accepted that they bind to poorly water soluble ligands in the cytosol, thereby facilitating improved intracellular solubilization[33], [34]. Several classes of iLBPs have been isolated, including sterol carrier proteins, cellular retinol binding proteins (CRABPs), fatty acid binding proteins (FABPs)[35] and BABPs, all of which display similar secondary and tertiary structures. A common β-clam shell fold, formed

by two orthogonal antiparallel five stranded β-sheets, is capped by a helix-turn-helix motif at the N-terminal domain, defining the so-called protein "open-end". The structures of apo and holo chicken L-BABP (cL-BABP) are reported in Figure 1.7.

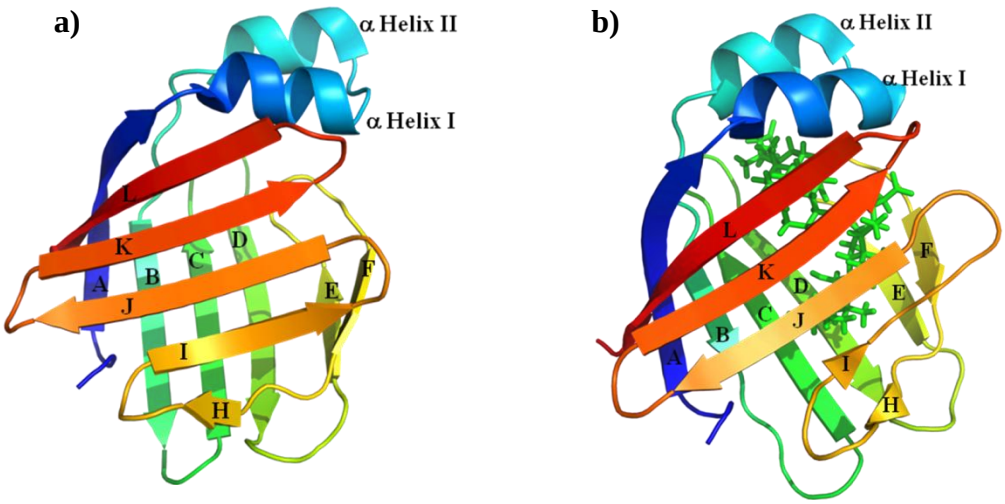


Figure 1.7 Three-dimensional structures of cL-BABP in its apo form[29] (PDB ID: 1MVG) (panel **a**) and in complex with two molecules of chenodeoxycholic acid[36] (PDB ID: 2JN3) (panel **b**).

BABPs are cytosolic lipid chaperones, which play central roles in driving bile flow, as well as in the adaptation to various pathological conditions, contributing to the maintenance of bile acid homeostasis and functional distribution within the cell. The understanding of the binding mode of bile acids with their cytoplasmic transporters is a key issue to provide a model for the mechanism of their transfer from cytoplasm to the nucleus.

Recent studies have shown that bile acids not only serve as the physiological detergents that facilitate absorption, transport, and distribution of lipid-soluble vitamins and dietary fats but also are the signaling molecules that activate nuclear receptors and regulate bile acid and cholesterol metabolism. In addition, bile acids induce the cytochrome P450 3A family of cytochrome P450 enzymes, to detoxify bile acids, drugs and xenobiotics in the liver and intestine, induce hepatocyte apoptosis and activate the gene encoding a candidate bile acid transporter protein[37]. Given the important role of bile acids, the study of their transport at a molecular level is of special medical and pharmacological interest. In this line it is essential to gain insight into the three-dimensional structures

and dynamic behavior of the proteins involved in bile acid recycling, both in their free and complexed forms.

1.4 Liver BABP and the role of disulphide bridges in intracellular proteins

The present thesis work deals with an NMR investigation of the role played by a naturally occurring disulphide bridge on the ligand binding, dynamics and functional properties of L-BABP. To this purpose we performed a comparative study on two liver BABPs, with and devoid of a disulphide bridge, both identified in chicken liver. The formation of S-S bridge is one of the factors, together with the chemistry of the ligand and the protein plasticity, that has been suggested to modulate bile salts selectivity, stoichiometry and binding affinity in BABPs[38].

To better understand the role of the disulphide bridge in the protein under study, a brief summary of the structural characteristics and functions of disulphide bonds, occurring in proteins located in the reducing cytoplasmic milieu, is given.

Disulphide bonds are known to play important role in proteins, ranging from folding to structure stability, to function. Not all the disulphide bonds in proteins are equally important and conserved. Indeed, small proteins with less than 70 residues, are often rich in disulphide bonds which are known to contribute critically to their stability[39], since they usually lack a strong hydrophobic core[40]. Proteins mainly characterized by α-helices show least preference for disulphide bridges, probably because of the steric hindrances involved in the formation of disulphides between cysteines present in different helices or in the same helix. The repertoire of cysteine conformations suggests preferences in disulphide distribution within β sheets[41] and highlights their preference for stabilizing strands and loops (Figure 1.8).

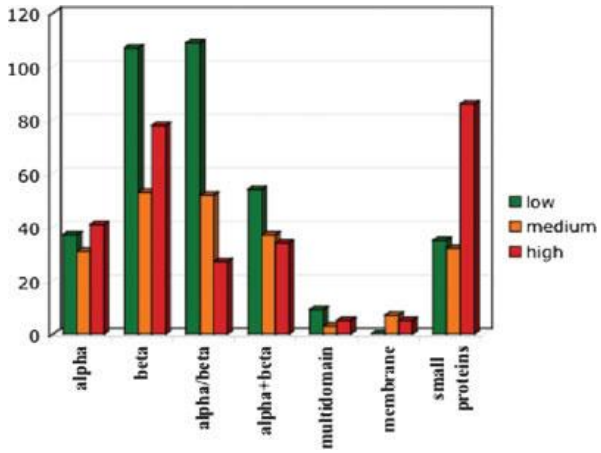


Figure 1.8 Distribution of high, medium, and low conserved disulphide bonds in SCOP (database containing disulphide bridge classification in protein families)[42].

Richardson and Thornton defined backbone protein contexts that could not form disulfides between resident Cys residues without introduction of strain into the protein structure. With respect to secondary structure, disulfide bonds should not be found between Cys pairs: (a) on adjacent β-strands[43], [44]; (b) in a single helix or strand[44]; or (c) on nonadjacent strands of the same β-sheet[44]. Whenever disulphide bonds are found in these contexts they are referred to as forbidden disulfides, recognized as Cross-Strand Disulfide (CSD). These CSD are expected to have a functional, redox-related role rather than a structural role.

An interesting example of a protein containing forbidden disulphide bridges, the so called β-flip, is the retinol binding protein[45], a carrier protein that, in complex with transthyretin, transports vitamin A within the blood stream to target cells.

Disulphide bonds are known to be present mainly in the extracellular proteins and rare in the cytoplasmic compartments of most organisms, due to the reductive nature of the cytosol[46], [47].

The chemical environment inside a typical cell is reducing, with a reduction potential around −200 to −300 mV for *Escherichia coli*. As a result, cysteine residues in intracellular proteins are generally found in their reduced form with free sulfhydryl groups.

However, although rare, there are examples of disulfide bonds in intracellular proteins. They are mainly found in proteins that catalyze oxidation-reduction (redox) processes. Among these there are members of thoiredoxin family, the thioredoxin and the glutathione reductase, in which a disulfide bond forms during the catalytic cycle. In

other systems such as Hsp33 and OxyR, the disulfide bond formation is part of a redox-sensing mechanism[48, 49], [50].

These functionally important disulphide bonds are usually very highly conserved even if the family members share very low sequence identity. They are however rare and generally transiently formed or marginally stable.

Several functional examples of disulphide bridges have been reported in the literature for members of the lipocalin family, where the introduction/removal of a disulphide bridge was responsible for changes in ligand binding stoichiometry and affinities.

Reduction of a disulfide bond in human tear lipocalin has been reported to play a relevant functional role as it induces a nine-fold increase in ligand affinity[51]. The removal of a disulfide bond in rat lipocalin-type prostaglandin D synthase slightly increased the binding affinity for biological ligands, by effect of allowing a less compact barrel pocket and a higher number of residues contributing to ligand binding[52]. In the cellular retinoic acid binding protein I the introduction of a disulphide bond made the binding of the retinoic acid irreversible, as a result of the quenching of portal region structural mobility[53]. Most of liver BABPs belonging to the non-mammalian species show a disulphide bond involving the conserved Cys 80 with the cysteine in position 91. The presence of a disulphide bridge in the protein scaffold of the homologous liver zebrafish BABP (69.8% of identity with chicken liver BABP, calculated with clustalW) was correlated with the exhibited binding stoichiometry[54] which varied from one to two ligand molecules, when the disulphide bridge Cys 80-Cys 91 was removed.

In the case of cL-BABP, the two cystein residues forming the naturally occurring disulphide bridge, namely Cys 80 and Cys 91, belong to two adjacent antiparallel β-strands, in conflict with Richardson and Thornton's rules. What is here expected, for cL-BABP, is that the disulfide bridge and the β-sheets around it are significantly distorted. For these reason the Cys pair in the protein under study does not seem to fit the paradigm of disulfides as structural stabilizer[32], [55], [29], [36], [30].

2. Material and Methods

2.1 Materials

2.1.1 Chemicals

All chemicals and reagents used were of analytical grade purity and, if not otherwise specified, were purchased from Sigma-Aldrich or Carlo Erba. ^{15}N enriched ammonium chloride, ^{13}C enriched glucose and D_2O were purchased from Cambridge Isotope Laboratories.

The DNA coding for cL-BABP, already cloned in the pET24d vector, has been originally provided by Dr. J. Foote (Fred Hutchinson Cancer Research, Seattle, USA).

pET24d(+) (Novagen) is a 5.3 Kb expression vector based on the strong T7 promoter. Unlike T5, T7 promoter is recognized only by T7 RNA polymerase, therefore T7 transcription/expression is induced by providing a source of T7 RNA polymerase in the host cell.

2.1.2 Culture media

Media for bacterial growth were prepared as follows, if not otherwise specified.

Luria-Bertani (LB) broth

For a final 1L:

- 10g triptone
- 5g yeast extract
- 5g NaCl

pH was adjusted to 7.0. After sterilisation, antibiotic was added.

Minimal medium (MM)

For a final 500 ml:

- 400 ml of sterile water

- 100 ml of M9 salts (vide infra)
- 1 ml of sterile 1 M $MgSO_4$
- 1 ml of sterile 50 mM $ZnSO_4$
- 500 µl of sterile 0.1 M $CaCl_2$
- 50 µl of 10 mM $FeCl_3$ (0.22 µl filtered)
- 10 ml of glucose from a 40% stock solution (0.22 µm filtered)
- 3 ml of vitamin stock solution (vide infra)

Sterilisation was achieved in autoclave if not otherwise specified. After sterilisation, antibiotic was added.

M9 salts

For a final 1L:

- 33.9 g Na_2HPO_4
- 15 g KH_2PO_4
- 2.5 g NaCl
- 5 g NH_4Cl

Sterilisation was achieved in autoclave.

Vitamin stock solution

For a final 200 ml:

- 100 mg thiamine
- 20 mg d-biotin
- 20 mg choline chloride
- 20 mg folic acid
- 20 mg niacinamide
- 20 mg d-panthothenate
- 20 mg pyridoxal hydrochloride
- 2 mg riboflavin

Sterilisation was achieved by 0.22 µm filtration.

In media containing bacteria transformed with pET24d kanamicine was used at a final concentration of 50 µg/ml.

2.2 Methods

Cloning and site directed mutagenesis procedures have been extensively described in Antonio Bertarelli Degree Thesis, titled: "Espressione e caratterizzazione di cL-BABP ed opportuni mutanti". Supervisors: Prof.ssa Henriette Molinari of the University of Verona and Dr Laura Ragona of ISMAC-CNR, Milano.

2.2.1 ^{15}N-^{13}C labelled cL-BABP/S-S (T91C cL-BABP) expression

A pre inoculate solution of 30 ml of cL-BABP/S-S in BL21 with 50 µg/ml Kan antibiotic was grown in M9 minimal medium, using $^{15}NH_4Cl$ and ^{13}C enriched glucose, shacking overnight at 37°C. BL21 (DE3) (Stratagene) is the *E. coli* strain for protein production in T7 promoter-driven vectors. The advantages of these cells are the high-level protein expression and the ease of induction.

The starting culture was used to inoculate 1L of M9, then incubated in a shaker at 37°C till an OD_{600} value of 0.6 was reached. Induction was achieved by addition 1 mM IPTG and incubation overnight at 22°C.

0.7 µl of β-ME per ml were added to the lysis buffer (vide infra) and PMSF (phenylmethanesulfonylfluoride) 1:2000. After a centrifugation step at 8000g for 10 minutes, the pellet was resuspended in 2.5 ml of lysis buffer + β-ME + PMSF/g of pellet. The buffer final volume was of 20.75 ml and sonicated with 4 cycles of 30 pulses at 60 Watt. The culture was then centrifuged at 10000g for 10 minutes. The pellet underwent another resuspension-sonication-centrifugation cycle, while the supernatant, containing the soluble protein fraction, was loaded on the first purification column.

Before induction a 1 ml sample was collected and centrifuged at 10000g for 5 minutes and the pellet, resuspended in 200 ml of sterile water, was kept for the evaluation of basal protein expression.

The protocol for the expression of ^{15}N labelled proteins or for unlabelled proteins was the same, except that a 30 ml starter culture and LB medium was used.

Lysis buffer:

- 50 mM Tris-HCl pH 8.0

- 20% sucrose
- 1 mM EDTA
- 1 mM PMSF freshly added

2.2.2 Protein purification (Purification protocols)

An anion exchange chromatography was the first step of purification for His-tagged recombinant cL-BABPs. Next steps were a size exclusion chromatography and a delipidation, followed by a buffer exchange in the storage/NMR buffer and liophylization.

2.2.2.1 Anion exchange chromatography

80 g of DE52 pre-swollen microgranular DEAE cellulose (Whatman) resin were resuspended in 200 ml of 0.2 M Tris-acetate pH 7.8. After sedimentation, the supernatant was removed and the resin washed twice with 200 ml of 50 mM Tris-acetate pH 7.8. The resin was then poured into a XK 26/40 (GE Healthcare) column and extensively washed with 50 mM Tris-acetate pH 7.8, 0.02% NaN_3 (running buffer) by using AKTAprime (GE Healthcare), a Fast Protein Purification Chromatography (FPLC) platform. The supernatant from the expression of the high pI recombinant protein was pumped onto the column at 1 ml/min. Elution of the protein of interest was achieved with several volumes of running buffer. Fractions of 10 ml were collected (see Figure 2.1) and analyzed by Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE), reported below in paragraph 2.2.3 (see Figure 2.2).
After use, the resin was regenerated according to the manufacturer's instructions.

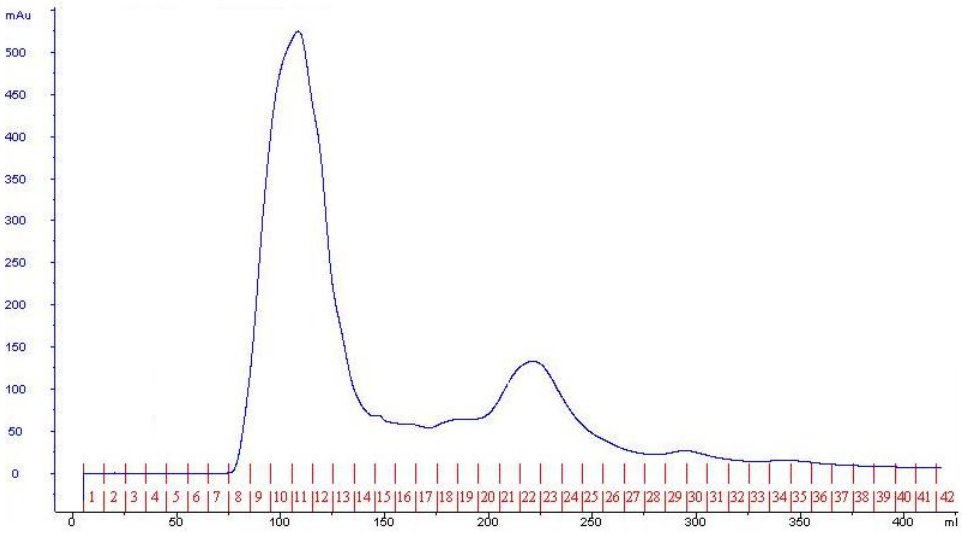


Figure 2.1. Elution profile obtained by the anion exchange column, increasing salt concentration. Eluted fractions, of 10 ml each, are reported in red along the *x*-axis. mAU: milli-absorbance unit.

Lane 1: marker
Lane 2: time zero
Lane 3: total fraction
Lane 4: soluble fraction
Lane 5: fraction number 8
Lane 6: fraction number 11
Lane 7: fraction number 14
Lane 8: fraction number 22
Lane 9: fraction number 26
Lane 10: fraction number 29

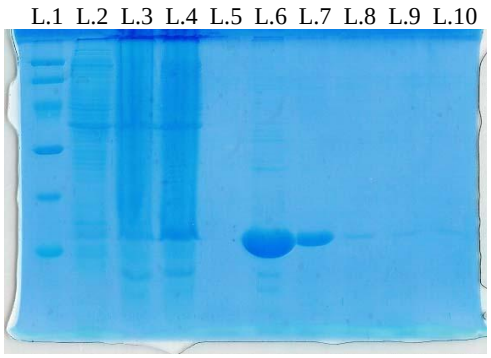


Figure 2.2 ^{15}N-^{13}C cL-BABP/S-S fraction poured on the SDS-PAGE gel (vide infra). Lane numbers, corresponding to the different eluted fractions, are reported at the top of the gel. Fractions containing the protein (corresponding to lanes 6 and 7) were concentrated with AMICON.

2.2.2.2 Size exclusion chromatography

500 ml of Sephacryl S-100 HR (GE Healthcare) resin were poured into a XK26/100 (GE Healthcare) column and extensively washed with 50 mM Tris-HCl pH 7.2, 0.2 M NaCl, 0.02% NaN_3 (running buffer) by using AKTAprime (GE Healthcare). Fractions containing the protein of interest, resolved with the previous chromatographic step, were concentrated to 5 ml and pumped onto the column at 1 ml/min (see Figure 2.3). Elution of the protein of interest was achieved with several volumes of running buffer. Fractions of 10 ml were collected and analyzed by SDS-PAGE (see Figure 2.4).

After use, the resin was regenerated according to the manufacturer's instructions.

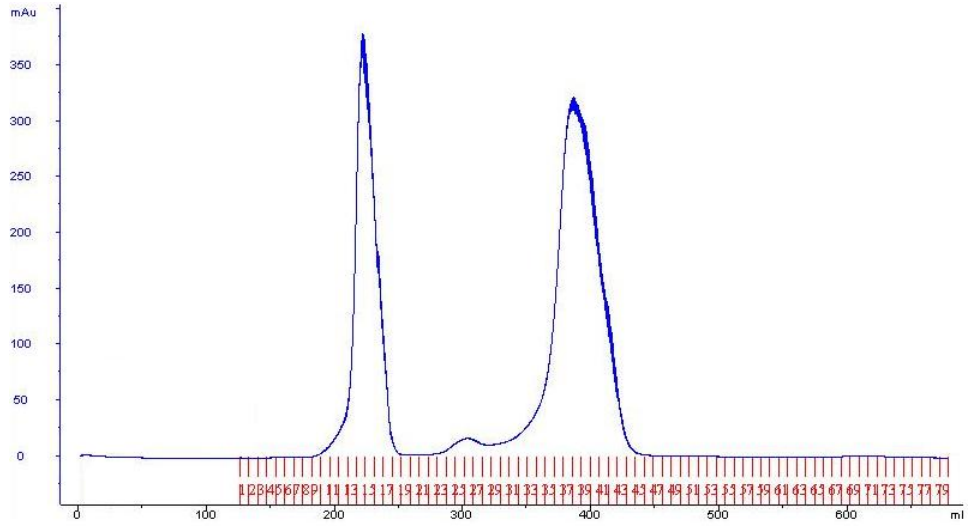


Figure 2.3. Elution profile obtained for concentrated fractions collected by the size exclusion column. Eluted fractions, of 10 ml each, are reported in red along the *x*-axis. mAU: milli-absorbance unit.

Lane 1: marker
Lane 2: fraction number 9
Lane 3: fraction number 14
Lane 4: fraction number 18
Lane 5: fraction number 26
Lane 6: fraction number 31
Lane 7: fraction number 38
Lane 8: fraction number 44
Lane 9: eluted from small amicon
Lane 10: eluted from big amicon

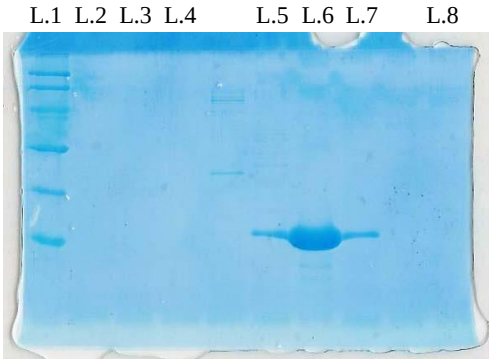


Figure 2.4 ^{15}N-^{13}C cL-BABP/S-S fraction poured on the SDS-PAGE gel (vide infra). Lane numbers, corresponding to the different eluted fractions, are reported at the top of the gel. Fractions containing the protein (corresponding to lanes 5, 6 and 7) were collected for the following delipidation.

2.2.2.3 Removing endogenous lipids

In order to eliminate endogenous ligands, a delipidation procedure[56] is required for proteins binding hydrophobic molecules expressed in *E. coli*. 30 ml of Lipidex 1000 (Packard Bioscience) resin were washed twice with 5 volumes of 10 mM sodium phosphate pH 7.4, 0.02% NaN_3 (running buffer, vide infra) at 4°C. The resin was then poured into a XK16/20 (GE Healthcare) column and extensively washed with the running buffer at a flow rate of 36 ml/hour. Prior to loading the sample, the resin was

equilibrated for an hour at 37°C by connecting the column to a preheated water bath. Fractions containing ^{15}N-^{13}C cL-BABP/S-S, resolved with the previous chromatographic step, were diluted at 1 mg/ml (after elution, the protein concentration was of 1.9 mg/ml for a total amount of 68 ml), and applied to the column using a peristaltic pump system at a flow rate of 12 ml/hour. Elution of the protein of interest was achieved at 37°C with several volumes of running buffer (vide infra) at a flow rate of 36 ml/hour. Fractions of 2 ml were collected and analyzed by SDS-PAGE (vide infra).
After delipidation, the protein was firstly washed with phosphate buffer 30 mM pH 7.0 + 250 mM NaCl + 0.02% NaN_3 and subsequently washed with phosphate buffer 30 mM pH 7.0+ 0.02% NaN_3.

2.2.3 SDS-PAGE

For protein analysis with Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS PAGE) a sample with protein (1-5 μg) was made up to 20 μl in reducing loading buffer (vide infra) and heated at 95°C for 5 minutes. Each 20 μl sample was loaded into a separate well of a discontinuous 15% polyacrylamide gel. A 5 μL sample of molecular weight marker (GE Healthcare) was also loaded in a separate well for internal referencing. Gels were run at 120 V for approximately one hour. On completion, the gel was immersed in the staining solution (vide infra) for 10 minutes, followed by 10 minutes in the destaining solution (vide infra) and overnight destaining in bidistilled water.

Reducing loading buffer:
- 50 mM Tris-HCl pH 6.8
- 100 mM DTT
- 2% SDS
- 10% glycerol
- 0.1% bromophenol blue
- Polyacrylamide gel
- Stacking gel solution

For a final 2 ml:

- 330 µl of 30% acrylamide mix (BioRad)
- 250 µl of 1 M Tris-HCl pH 6.8
- 20 µl of 10% SDS
- 20 µl of 10% ammonium persulfate
- 2 µl of TEMED

Running gel solution

For a final 4 ml at 15% of acrylamide:

- 2 ml of 30% acrylamide mix (BioRad)
- 1 ml of 1.5 M Tris-HCl pH 8.8
- 40 µl of 10% SDS
- 40 µl of 10% ammonium persulfate
- 16 µl of TEMED

Running buffer:

- 25 mM Tris-HCl pH 8.3
- 250 mM glycine
- 0.1% SDS

Staining solution (for a final 1L):

- 400 ml of methanol
- 100 ml of acetic acid
- 1.5 g Comassie Brilliant Blue

Destaining solution (for a final 1L):

- 50 ml of methanol
- 75 ml of acetic acid
- 20 ml of glycerol

2.2.4 Estimation of protein concentration

Protein concentration estimates (c) were calculated from the Beer-Lambert equation:

$$Abs_{280nm} = \varepsilon_{280nm} \times l \times c$$

The absorbance at a wavelength of 280 nm (Abs_{280nm}) was detected using a Unicam UV2 UV/VIS spectrometer. The path length (l) of the sample in a measurement cuvette was 1 cm. A theoretical extinction coefficient at a wavelength of 280 nm measured in water (ε_{280nm}) was estimated for each construct using the ProtParam Tool from the Expasy Tools server (http://www.expasy.ch/tools/protparam.html).

2.2.5 Sample storage

Protein samples were exchanged into a buffer suitable for NMR experiments and concentrated to 0.5 mM. to minimize protein aggregation and precipitation.

Protein Buffers for:
cL-BABP: 30 mM potassium phosphate pH 7.2 or 6.5, 135 mM KCl, 10 mM NaCl, 1 mM DTT, 0.05% NaN_3
cL-BABP/S-S: 30 mM potassium phosphate pH 7.2 or 6.5, 135 mM KCl, 10 mM NaCl, 0.05% NaN_3. In this case DTT has not been used to avoid disulphide bridge reduction.
Samples were then lyophilized and stored at -20°C.

2.3 NMR analysis

2.3.1 Sample preparation

For NMR analysis lyophilized samples were resuspended in 90% double distilled water/10% D_2O (v/v) or 100% D_2O (v/v).
In particular L-BABP samples of 4.2 mg were resuspended in 540 µl of H_2O + 60 µl of D_2O, while samples of 2.8 mg were resuspended in 360 µl of H_2O + 40 µl of D_2O.

Unenriched bile acids and [24-^{13}C]-glycocholate (GCA) were purchased from Sigma. ^{15}N glycine conjugates of of chenodeoxycholate (CDA) and cholate (CA) were prepared as follows: ^{15}N Fmoc-Glycine was anchored to 2-chlorotrityl chloride resin in the presence of N,N-diisopropylethylamine according to the procedures of Barlos et al[57]. After removal of the Fmoc group by 20% piperidine in dimethylformamide and further washings, a pre-activated pentafluorophenyl ester of CDA or CA in dimethylformamide was added in the presence of 1-hydroxybenzotriazole to ^{15}N-Glycine anchored to the resin and vortexed until the ‘‘Kaiser test’’16 was negative. After extensive washings to remove the reagents, GCDA and GCA ^{15}N enriched in the glycine moiety were removed from the resin by repeated treatment with 25% 1,1,1,3,3,3-hexafluoro-2-propanol in dichloromethane. Solvents were removed in vacuum, the adduct was dissolved in aqueous acetonitrile and concentrated, with concomitant precipitation of the products, recovered by filtration. The homogeneity of the derivatives was assessed by Thin Layer Chromatography.

The titration of the unlabelled cL-BABP/S-S protein with increasing amounts of ^{15}N GCDA or ^{15}N GCA has been performed at 7 protein:ligand ratios (1:0.3, 0.6, 1, 1.5, 2, 2.5, 3). To prepare the holo-protein samples, bile salts were dissolved in tetrahydrofuran, and the concentration of the stock solutions was determined by measuring dry weights. Appropriate amounts of the stock solution of bile acid were aliquoted, and the solvent was evaporated under a stream of nitrogen. The bile acid was solubilized using 1:1 equivalents of 1M NaOH. Then 17.5 μL of D_2O were added to 332.5 μL of protein solution to bring the total volume of the NMR sample to 350 μL.

To acquire these experiments “shigemi” NMR tubes were used. The final 0.5 mM protein samples were dissolved in 30 mM potassium phosphate buffer, in 95%H_2O/5%D_2O. The pH of the solutions was 7.2. To minimize errors each ligand:protein mole ratio sample was prepared and analyzed twice.

For protein-ligand titration experiments a ^{15}N cL-BABP/S-S stock solution concentration was diluted to 0,2 mM. The concentration of the ligand stock solutions were determined by measuring dry weights using a microbalance.

Glycochenodeoxycholic acid was dissolved in the same 30 mM phosphate buffer (pH 7.2). The ligand was added to the protein in small amounts to achieve saturation within 10 steps (P:L ratios employed 1:0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.3, 1.6, 2.0), and three

additional aliquots were added to ensure the end of titration (P:L ratios 1:2.5, 3.0, 3.5). To acquire these experiments typical 5mm diameter NMR tubes were used.

2.3.2 Data collection

The majority of NMR experiments, acquired for assignments, were recorded on a Bruker DMX 500 spectrometer, operating at 500.13 MHz proton Larmor frequency, equipped with a triple resonance probehead at the NMR laboratory of ISMAC-CNR, Milano, or on a Bruker Avance III 600 spectrometer, operating at 600.13 MHz proton Larmor frequency, equipped with a triple resonance TCI cryoprobe, incorporating gradients on the z-axis, at the NMR laboratory of the University of Verona.
The NMR parameters used for setting the experiments are reported in the "Experimental Procedure" section of Cogliati, C., et al, *FEBS Journal*, 2009[38], inserted in Chapter 6.
Spectra for the titrations of cL-BABP and cL-BABP/S-S with GCDA, were recorded on a Bruker AVANCE 500 MHz spectrometer equipped with a cryogenically cooled 5mm TXI probe, at the HWB-NMR Facility (University of Birmingham, UK).
Relaxation Dispersion experiments were acquired on Varian Inova spectrometers equipped with room temperature probe heads at two static magnetic fields, corresponding to proton Larmor frequencies of 600 and 900 MHz, at the HWB-NMR Facility (University of Birmingham, UK).
The NMR parameters used for setting cL-BABPs titration and RD experiments are reported in the "Experimental section" of the manuscript "Site-specific investigation of steady-state kinetics and dynamics of the multistep binding of bile acid molecules to a lipid carrier protein.", submitted to *Chemistry. A European Journal* and inserted in Chapter 7
Experiments requiring a higher sensitivity were recorded on a 4-channel Varian Inova operating at 799.766 MHz equipped with HCN 5mm z-PFG cryogenic probe with enhanced ^{13}C and ^{1}H sensitivity and on a Varian Inova spectrometer equipped with a room temperature probe head with a corresponding Larmor frequency of 600 MHz, at the HWB-NMR Facility (University of Birmingham, UK). The experimental temperature was set to 25 °C unless otherwise specified.

2.3.3 Data acquisition, processing and visualization

The software Topspin (Bruker, headquartered in Billerica, Massachusetts, USA) and NMRPipe[58] were used for data acquisition and processing on Bruker instruments while VnmrJ (Varian, headquartered in Palo Alto, CA, USA) and NMRPipe were used for data acquisition and processing on the Varian spectrometers.

NMRLAB toolbox was used for NMR data processing in MATLAB (The Mathworks, headquartered in Natick, Massachusetts, USA). MATLAB is a matrix-oriented high-level programming environment which gives access to fast algorithms for a large number of numerical tasks on many common computer platforms UNIX, Windows. Data processing can be achieved either by scripts or by a user-friendly command structure. The advantage in using NMRLab is that the output files, obtained by the processing of Bruker or Varian data, are MATLAB files and can be directly used in NMRKIN software without any previous conversion in ASCII or MATLAB versions[59].

3. NMR for protein resonance assignments

NMR technique, when applied to the study of biological macromolecules in solution, allows the determination of their tridimensional structures, together with the characterisation of dynamics and protein-ligand (putative ligands, small organic molecules, drugs, nucleic acid), protein-protein interactions.

NMR has evolved considerably during the last years: high-field spectrometers and ultra-sensitive cryoprobes allow completely new experiments and applications, and new techniques have been developed to study molecules up to sizes as large as 100 kDa.

This chapter deals with an overview of the NMR experiments used in my research. Standard 2D and 3D experiments will be briefly described, while a considerable amount of details will be given on the analysis of 3D triple resonance experiments. If more than two nuclei are involved, letters like H, N and C will represent the ^{1}H, ^{15}N and ^{13}C magnetization, respectively. The angular momentum operators Ix, Iy and Iz represent the x, y, z components of the spin angular momentum of the system.

3.1 2D NMR

In a 1D spectrum the signal is recorded as a function of one time-variable. Therefore, after Fourier transformation, a spectrum displaying intensities versus frequencies is obtained. In this spectrum, each peak will correspond to one frequency. If two or more spins are scalar-coupled, the fine structure of a multiplet will be observed. In 1D spectra, of complicated molecules like proteins it is difficult to establish which spins are coupled to each other.

In a 2D spectrum the signal is detected as a function of two time-variables t_1 and t_2. Instead of recording one spectrum, a series of spectra is recorded where the time t_1, which corresponds to the delay between two RF pulses, is systematically incremented. Therefore, the recorded signal is a matrix, with axes corresponding to t_1 and t_2. Fourier transformation has to be performed in both t_1 and t_2 directions, and each signal will be a function of this two frequency variables.

The scheme below represents the general idea of a 2D experiment:

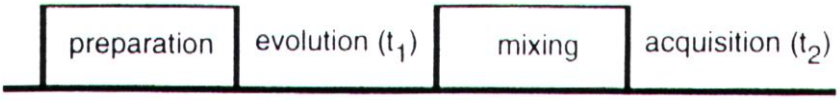


Figure 3.1 General scheme for a 2D experiment[60],[61].

The first period called preparation period, usually corresponds to the relaxation delay, which is used to restore all the magnetization along the z-axis (equilibrium state). Then one pulse creates transverse magnetization which is allowed to evolve during t_1. During the mixing period there are different pulses that end up converting one coherence into another. At the end of the mixing period the transverse magnetization is detected as function of the second variable, t_2, the acquisition time. The t_1 period is also called evolution period, because during t_1 the chemical shift of the desired nucleus evolve and, as a result, the observed ^{1}H signal will be modulated by the frequency of the hydrogen, or nitrogen, or carbon scalarly or dipolarly coupled to it, depending on the experiment we are acquiring. The resulting 2D spectrum will display a cross peak between connected resonances. Therefore the main difference between 2D and 1D NMR is that in the first it is possible to know which spins are coupled to which. Moreover in 2D experiments the overlapping of resonances of the 1D spectrum is reduced because the information is spread in the second dimension.
Generally unwanted coherence transfers can be eliminated by using phase-cycling, field gradient pulses or by a combination of both.

3.1.1 TOCSY

The TOCSY experiment[62], is used to observe connectivities between the protons within the same amino acid. As in the COSY experiment[60], the magnetization is transferred through scalar coupling. The advantages of the TOCSY over the COSY experiment is a higher sensitivity for larger molecules, all absorptive line-shapes, and the fact that correlations between all nuclei along a spin system are obtained.

The pulse sequence is shown in figure 2.2:

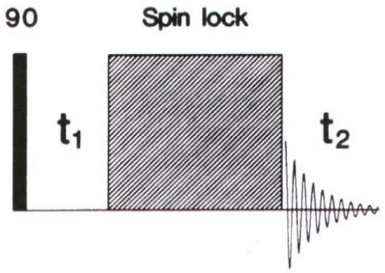


Figure 3.2 Basic pulse sequence of TOCSY experiment. A first 90° excitation pulse is followed by a variable delay time t_1, a spin-lock pulse and the recording of the FID.

The transfer of in-phase coherence from one spin to another is obtained using a spin-lock period applied in the mixing step. The spin lock is a so-called train of non-selective 180° pulses. It eliminates the chemical shift evolution, and only *J*-modulation evolves. The result is that relayed and multiple-relayed peaks are observed.

3.1.2 NOESY

The NOESY[63] spectrum is the most important source of structural information of a molecule. The basic idea is to create connectivities between the spins via cross-relaxation.

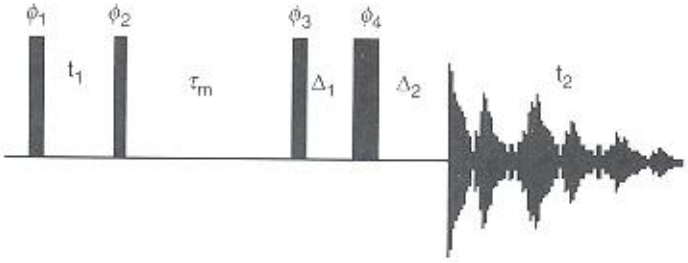


Figure 3.3 Basic pulse sequence of a 2D NOESY [64]. Thin and thick bars indicate nonselective 90° and 180° pulses, respectively; t_1 is the evolution time. τ_m is the mixing time[65].

The experiment (Figure 3.3) starts with a relaxation delay to allow the Z-magnetization to be restored after a previous scan, and a 90° pulse to create transverse magnetization (Iz→Iy and Sz→Sy). In the evolution step, t_1, the spins are allowed to precess so that the terms Ix, Iy, Sx and Sy are created and labelled with their own frequencies. Here,

the *J*-coupling between them is neglected. The second 90° pulse brings the magnetization back to the z-axis. During the mixing time τ_{mix}, the system undergoes relaxation and longitudinal magnetization is transferred from one spin to another, Iz → Sz, via cross relaxation. The final 90° pulse brings the magnetization to the xy-plane where the Larmor precession will be monitored during t_2. The intensity of the NOESY cross-peak is proportional to the cross relaxation rate between the two spins I and S, which is dependent on their internuclear distance.

The relation between inter-nuclear distance and peak intensity is given by

$$NOE = C/r^6 \qquad (1)$$

where to obtain the constant C, a previous calibration with a known distance, or distances, is necessary and r is the distance between the nuclei giving rise to the cross peak.

The structural information is extracted from NOESY spectra comparing the intensity of each assigned cross-peak with the one taken as reference, between two protons whose distance is known and constant, such as, for example, the two protons Hδ and Hε of a tyrosine aromatic ring, or the two methylenic protons.

3.1.3 HSQC

The ^{1}H-^{15}N-HSQC and ^{1}H-^{13}C-HSQC experiments are the basic ones to test protein samples and their interactions with ligands and also form the basis of heteronuclear 3D experiments. The ^{1}H-^{15}N-HSQC displays cross-peaks for each proton bound to nitrogen, in particular backbone or side chain amide protons, while the ^{1}H-^{13}C-HSQC displays cross-peaks for each proton attached to carbon.

The ^{1}H-^{15}N-HSQC spectrum is considered a “fingerprint” of the protein. The dispersion of the cross-peaks indicates whether the protein is folded and the feasibility and ease of the resonance assignment. It is also used to measure the rates of proton-deuterium exchange of the labile protons and in titration experiments of proteins with different ligands to identify the residues that are involved in interactions.

The pulse sequence for the ^{1}H-^{13}C or the ^{1}H-^{15}N-HSQC experiment is shown in Figure 3.4.

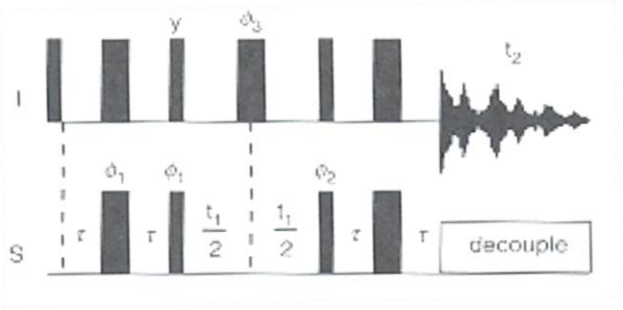


Figure 3.4 HSQC experiment pulse sequence. I = ^{1}H, S = ^{13}C or ^{15}N. The pulses in the 1H channel (which affect only the protons) and those in the heteronucleus channel are shown separately. Thin and thick bars indicate nonselective 90° and 180° pulses, respectively. The 90° pulse followed by a relaxation delay, τ, a 180° pulse, another relaxation delay and finally by a 90° pulse represents an INEPT block[65].

A first inspection of this pulse sequence shows a basic building block of pulse programs called INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) [66], followed by an evolution period (t_1), another INEPT scheme and then detection. The first INEPT is used to transfer magnetization from the ^{1}H to the heteronucleus (S = ^{15}N or ^{13}C). The t_1 period that follows the first INEPT is incremented to build up the indirect ^{15}N or ^{13}C dimension. The second INEPT brings the magnetization back to proton, the observed channel. The pathway of the magnetization flow during this experiment is represented below:

$$H^S \xrightarrow{J_{HS}} {}^S S\,(t_1) \xrightarrow{J_{SH}} H^S\,(acq)$$

The cross peaks don't show fine structures due to the NH or CH coupling in the indirect and direct dimensions, because during t_1 the evolution of the *J*-coupling is refocused by the 180° pulse applied on ^{1}H in the middle of the evolution period, and during acquisition, a broadband decoupling is applied.

3.2 3D NMR

In a 3D experiment all the concepts of 2D NMR are employed. The 3D experiments are essential to study big systems such as proteins, because the overlap observed in 2D spectra can be drastically reduced by extending the information to a third dimension. The 3D NMR pulse sequences are derived from a combination of two 2D experiments as shown below (Figure 3.5).

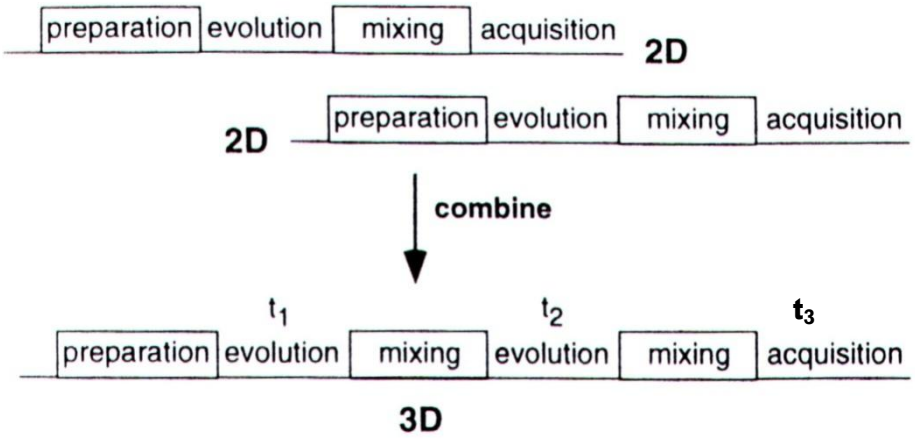


Figure 3.5 Schematic generation of a 3D NMR experiment from the combination of two 2D experiments. The mixing period of one 2D experiment and the preparation period of a second one are combined. The 3D experiments contain three independent time periods. The FID is acquired during the acquisition period, t_3, as a function of two independently incremented evolution times, t_1 and t_2. A mixing period follows each evolution time[65].

The third dimension is created by introducing a second evolution period, t_2. Therefore, in a 3D experiment the signal, which is monitored during the acquisition time t_3, is function of two evolution times, t_1 and t_2, which are incremented independently. The signal is a matrix with three axes corresponding to t_1, t_2 and t_3. After Fourier transformation a spectrum with three different frequency dimensions F1, F2, F3 is obtained. This spectrum can be imagined as a cube (Figure 3.6).

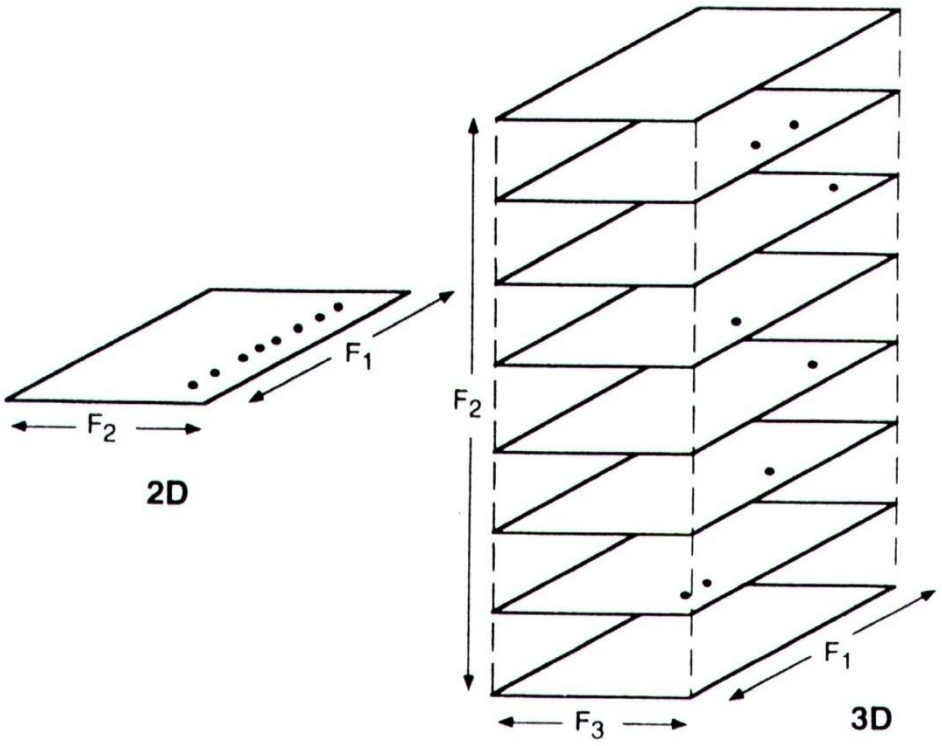


Figure 3.6 The development of a 3D data set from a 2D data set. The 2D depicted here shows a set of resonances that, although resolved in F_1 dimension , are not clearly determined in the F2 dimension. The introduction of an additional evolution period generates a third frequency dimension perpendicolar to the first two. The increased resolution afforded by virtue of a second magnetization transfer step can alleviate ambiguities in the 2D spectrum[67].

3.2.1 ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC

3D ^{13}C or ^{15}N edited NOESY experiments[68], [69], [70] are important in protein NMR. They reduce the overlap of the crowded 2D-NOESY spectra by introducing a ^{13}C or ^{15}N dimension, so that each ^{15}N or ^{13}C plane of the 3D spectrum corresponds to a normal 2D NOESY spectrum.

From the ^{15}N-NOESY-HSQC spectrum, short-range NH-NH, intra-residue, sequential and long-range NOEs are obtained. In the ^{13}C-NOESY-HSQC spectrum, correlations within and between side-chains, long-range NOEs, including those involving aromatic protons are obtained. With such a big amount of information, these experiments are useful not only for the assignment of the backbone, the side chains resonances but also to collect NOEs for structural calculation.

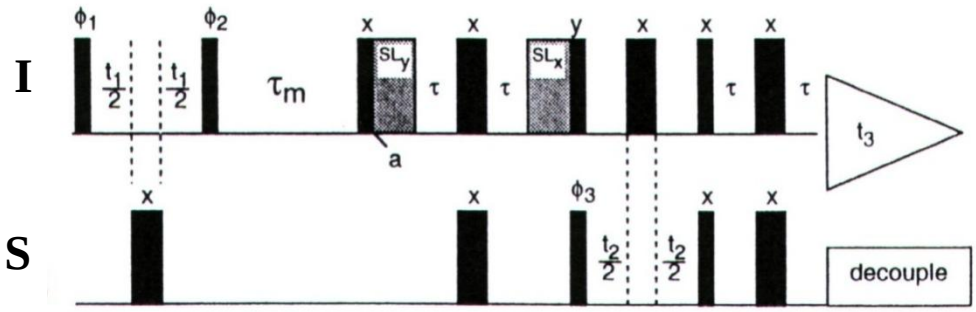


Figure 3.7 Basic pulse sequence of NOESY-HSQC experiment. Thin bars represent 90° pulses and thick bars represent 180° pulses. Pulses are applied with *x*-phase unless the phase is indicated above the bar. The nominal value for $2\tau = 1/(2J_{IS})$ and τ_m is the mixing time. In this scheme I = ^{1}H and S = ^{15}N or ^{13}C. SL indicated the Spin-Lock[65].

The first part of the pulse sequence (Figure 3.7) corresponds to a normal NOESY experiment, except for the 180° pulse applied on the ^{13}C or ^{15}N channel during the t_1 evolution period. This pulse is used to decouple $^{1}J_{IS}$ evolution. After that there is a HSQC block, where magnetization is transferred to ^{13}C or ^{15}N, with chemical shift evolution during t_2. Finally, magnetization is transferred back to protons and prepared for acquisition.

3.2.2 ^{15}N-TOCSY-HSQC

The ^{15}N-TOCSY-HSQC[69] is used to assign the side chains by correlating the intra-residue N^H, ^{15}N, and the side-chain protons using TOCSY transfer schemes. It is usually sufficient to assign the side chains of proteins up to 10 kDa. Unfortunately, in larger proteins, the decay of proton magnetization is usually very fast and competes with the coherence transfers involving small proton-proton *J*-couplings, thus making the ^{15}N-TOCSY-HSQC less efficient. However, it is still useful to correlate the intra-residue HN, ^{15}N and H^{α}, being able, in several cases, to go even further in the side chain.

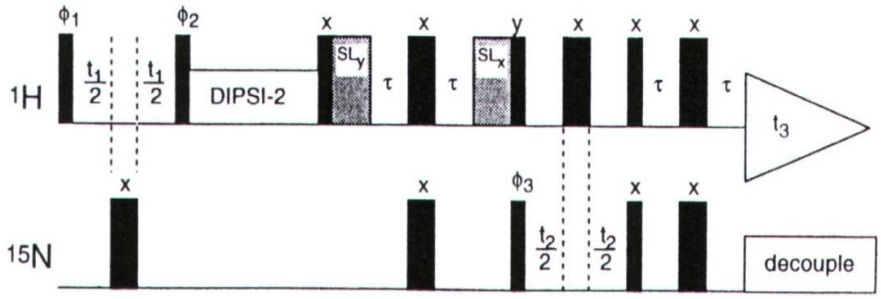


Figure 3.8 Basic pulse sequence for a ^{15}N- TOCSY-HSQC. Thin bars represent 90° pulses and thick bars represent 180° pulses. Pulses are applied with *x*-phase unless the phase is indicated above the bar. The nominal value for $2\tau = 1/(2J_{IS})$ and τ_m is the mixing time. DIPSI-2 isotropic mixing used in τ_m. SL indicated the Spin-Lock [65].

Like the ^{15}N-NOESY-HSQC experiment, the ^{15}N-TOCSY-HSQC can be described using two main blocks (Figure 3.8): the first block corresponds to a general TOCSY experiment (with 180° pulse for decoupling $^1J_{HN}$ evolution during t_1) and the second corresponds to a HSQC experiment. A "clean" version of the ^{15}N-TOCSY-HSQC has been here employed as the artefacts due to ROESY effects have less probability to occur[71]. In this experiment, the conventional 180° pulse train is replaced by series of 180°-90°-180° pulses separated by a certain delay Δ. By choosing an adequate value for Δ the NOE and ROE effects cancel each other, so that the magnetization is transferred only *via J* coupling.

3.2.3 HCCH-TOCSY

For large proteins (generally >10 kDa) the proton line widths are often too large and hinder efficient coherence transfers involving the relatively small ^{1}H-^{1}H *J* couplings, making the homonuclear or ^{15}N edited TOCSY experiments very inefficient. Two 3D HCCH-TOCSY experiments were introduced, one aimed at the assignment of carbon resonances and one used for assigning proton resonances of the side chains of ^{13}C- or (^{15}N,^{13}C)-labelled proteins[72],[73]. A general pulse sequence scheme for the H(C)CH-TOCSY experiment is given in figure 3.9.

The HCCH-TOCSY experiments rely on magnetization transfer between ^{13}C nuclei by isotropic mixing. Since the coherence transfer can be carried out in shorter periods of time due to the larger ^{13}C-^{13}C *J* couplings (> 35Hz), the experiments are well applicable

to larger molecules. Eventually, magnetization is transferred to protons using the large heteronuclear ^{1}H-^{13}C *J*-coupling. Therefore, the limiting factor of the HCCH *J*-correlation methods is the ^{13}C line width.

In the 3D (H)CCH-TOCSY spectrum two of the three dimensions correspond to ^{13}C and the other one to ^{1}H. Using this spectrum it is possible to assign all the carbons of the side chains. With the 3D H(C)CH-TOCSY spectrum, in which two of the three axes correspond to protons and the third to carbon, side chain protons are assigned.

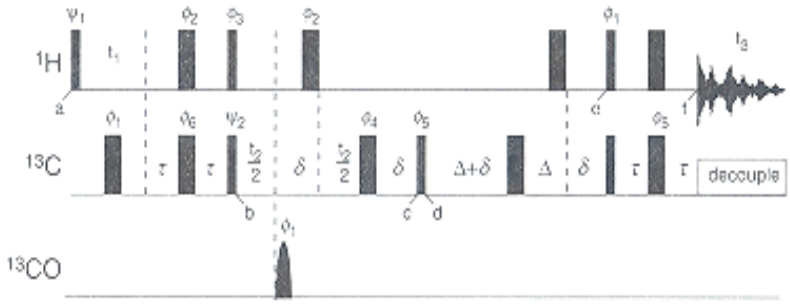


Figure 3.9 Pulse sequence of H(C)CH-TOCSY. Thin bars represent 90° pulses and thick bars represent 180° pulses. Pulses are applied with *x*-phase unless the phase is indicated above the bar. Rounded bars represented selective 180° pulses applied to ^{13}CO spin[65].

The H(C)CH experiment starts with a 90° pulse on protons, followed by evolution of the proton spins during the t_1 period. In the centre of t_1 a 180° pulse is applied on ^{13}C in order to refocus the $^{1}J_{HC}$ couplings. Afterwards, one INEPT block is applied to transfer magnetization from ^{1}H to ^{13}C ($H_y \rightarrow H_zC_y$). The duration is 2τ and the optimal value is $t=1/4\ J_{CH}$. In the next block, of total duration $2\delta+t_2$, the carbon magnetization that is in antiphase with respect to its directly attached proton spin is refocused, and the ^{13}C magnetization evolves during t_2. The delay δ has to be chosen to allow transfer of CH, CH_2 and CH_3, generally $\delta = 1/6\ J_{CH}$. A 180° pulse applied on the carbonyl refocus the $J_{C\alpha C'}$ coupling during this period. The next step of the sequence is the $^{13}C \rightarrow {}^{13}C$ transfer step. A TOCSY-like mixing unit, often a DIPSI composite pulse, is used to transfer in-phase ^{13}C magnetization to its neighbours. The spin lock in front of the DIPSI is used to destroy all in-phase carbon coherences that are not aligned with the spin lock field.

During the DIPSI period there is in-phase transfer among the carbons spins not only to direct bound carbons. The final step is an INEPT block that brings back the

magnetization to the attached protons which are detected during t_3. The flow of the coherence transfers in this experiment is given by:

$$H_i\,(t1) \xrightarrow{{}^1J_{CH}} {}^{13}C_i\,(t2) \xrightarrow{J_{CC}} {}^{13}C_j \xrightarrow{{}^1J_{CH}} H_j\,(t3)$$

The pulse sequence for the (H)CCH-TOCSY experiment can be derived from the previous one. The main difference is that the t_1 evolution period is now used to frequency label the ^{13}C spins, instead of the proton spins. The flow of magnetization transfer during the (H)CCH-TOCSY experiment is given by:

$$H_i \xrightarrow{{}^1J_{CH}} {}^{13}C_i\,(t1) \xrightarrow{J_{CC}} {}^{13}C_j\,(t2) \xrightarrow{{}^1J_{CH}} H_j\,(t3)$$

3.3 Triple resonance experiments

An alternative assignment run on a (^{13}C,^{15}N) doubly-labelled sample and is very useful for proteins with MW in the range 10-30 kDa, as in our case. This strategy involves 3D triple resonance experiments that correlate backbone H^N, $^{15}N^H$, H^α, $^{13}C^\alpha$, ^{13}CO, and side chain H^β and $^{13}C^\beta$ spins using one-bond heteronuclear coupling constants[74]. In this way the magnetization moves across peptide bonds, providing information that can be used for sequential backbone assignment. In the following paragraphs an overview of the triple resonance experiments performed in this study is given.

3.3.1 HNCO and HNCA

The HNCO and the HNCA are the basic 3D triple resonance experiments. In the first one the chemical shifts of the H^N proton and the ^{15}N of one residue (i) are correlated with the ^{13}CO of the previous residue (i-1), while in the second one the cross-peaks correlate each H^N with the $^{13}C^\alpha$ of residue i and i-1. The pulse sequence for the HNCO experiment is schematized in Figure 3.10.

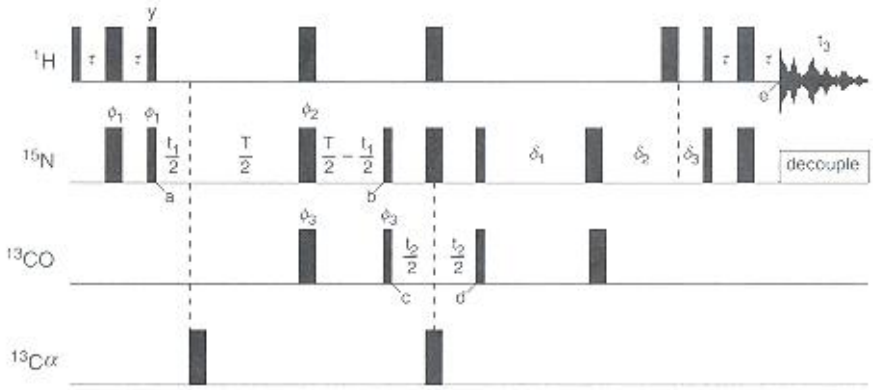


Figure 3.10 Pulse sequence of HNCO. Thin bars represent 90° pulses and thick bars represent 180° pulses. Pulses are applied with *x*-phase unless the phase is indicated above the bar[65].

The HNCO pulse sequence starts with the transfer of H^N magnetization to ^{15}N via one INEPT block. The next step is a constant-time evolution period of the ^{15}N-chemical shift (t_1). During this period, J_{HN} evolves for a time 2δ, while J_{NCO} evolves during 2Δ. By setting δ to $1/4J_{HN}$ (2.75 ms) and Δ (<16.7 ms) to $1/4J_{NCO}$ one obtains refocusing of HzNy antiphase magnetization and creation of NyCOz anti-phase magnetization. A 180° pulse is applied on $^{13}C^{\alpha}$ to refocus $J_{NC\alpha}$ modulation. After t_1, ^{15}N anti-phase magnetization with respect to ^{13}CO is converted into ^{13}CO anti-phase magnetization with respect to ^{15}N by a pair of 90° pulses applied on both channels. During t_2 there is evolution of the ^{13}CO chemical shift, while the *J*-modulations due to J_{NCO}, $J_{COC\alpha}$ and $J_{H\alpha CO}$ couplings are refocused by 180° pulses. After the t_2 period magnetization is transferred back to nitrogen by a pair of 90° pulses, after that NyCz is refocused to Nx during the delay 2Δ, and Nx converted into HzNy during the delay 2δ. A final pair of 90° pulses on both ^{15}N and ^{1}H channels transfer magnetization from ^{15}N to H^N. During the final INEPT this coherence is refocused to observable HN-magnetization. In summary, the overall flow of magnetization during the HNCO pulse sequence is given by:

$$HN \xrightarrow{J_{HN}} {}^{15}NH\ (t1) \xrightarrow{J_{NCO}} {}^{13}CO\ (t2) \xrightarrow{J_{NCO}} {}^{15}NH \xrightarrow{J_{HN}} HN\ (t3)$$

Note that the $^{13}C^{\alpha}$ and the carbonyl ^{13}CO are pulsed separately and independently. The ^{13}C-channel is split into two bands, one around 55 ppm, where the C^{α} carbons resonate, and another around 175 ppm where the carbonyl atoms resonate. This is performed by either switching the RF carrier frequency of the ^{13}C-transmitter to the appropriate

frequency , or by phase-modulating the ^{13}C-transmitter source with an appropriate offset frequency.

The pulse sequence for the HNCA experiment is the same as the HNCO, except by the fact that the t_2 evolution period is on $^{13}C^{\alpha}$ instead of ^{13}CO.

Here, as in the HNCO, the magnetization starts on the amide proton, and is transferred to $^{15}N^{H}$ by the first INEPT block. The evolution of the ^{15}N chemical shift is performed using a constant-time period. The one-bond $^{1}J_{NC\alpha}$ intra-residue coupling (8-12 Hz) and two-bond $^{2}J_{NC\alpha}$ inter-residue coupling (up to 7 Hz) will evolve simultaneously during t_1. As a result, the magnetization is transferred to both the intra-residue (i) and the previous residue (i-1) $^{13}C^{\alpha}$ spins. After t_1, the ^{15}N antiphase magnetization is converted into $^{13}C^{\alpha}$ antiphase magnetization by two simultaneous 90° pulses on ^{15}N and $^{13}C^{\alpha}$, and the $^{13}C^{\alpha}$ chemical shift evolves during t_2. During this period, the evolution of $^{1}J_{HC\alpha}$, $^{1}J_{NC\alpha}$ and $^{1}J_{COC\alpha}$ scalar couplings is refocused by 180° pulses. However, the scalar coupling between the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ spins is still active and therefore the acquisition time in this dimension should be shorter than $½J_{C\alpha C\beta}$. Since the T_2 relaxation time of C^{α} (10-20 ms for large proteins) is comparable to $½J_{C\alpha C\beta}$ = ½. 35Hz=14 ms, usually the acquisition times in the Cα dimension are not larger than 10-12 ms, such that $2J_{C\alpha C\beta}$ is not influencing the line width. $2J_{C\alpha C\beta}$ can be decoupled by a selective 180°pulse on C^{α} but not for all spin systems. Homonuclear $2J_{C\alpha C\beta}$ can also be refocused by a constant time C^{α} evolution period.

Finally, the magnetization is transferred back to the amide nitrogen, and converted into observable H^{N}-magnetization by the reverse pathway. The overall flow of the H^{N} magnetization during the HNCA experiment is given by:

$$HN \xrightarrow{J_{HN}} {}^{15}NH\ (t1) \xrightarrow{{}^{1}J_{NC\alpha}} {}^{13}C\alpha\ (t2) \xrightarrow{{}^{1}J_{NC\alpha}} {}^{15}NH \xrightarrow{J_{HN}} HN\ (t3)$$

3.3.2 CBCA(CO)NH

The CBCA(CO)NH[75] generates correlations between N_i-N^{H}_{i}-C^{α}_{i-1}-C^{β}_{i-1}. Besides that, the assignment of C^{α} and C^{β} chemical shifts also helps in identifying the amino acid. A brief analysis of the CBCA(CO)NH pulse sequence is given in Figure 3.11.

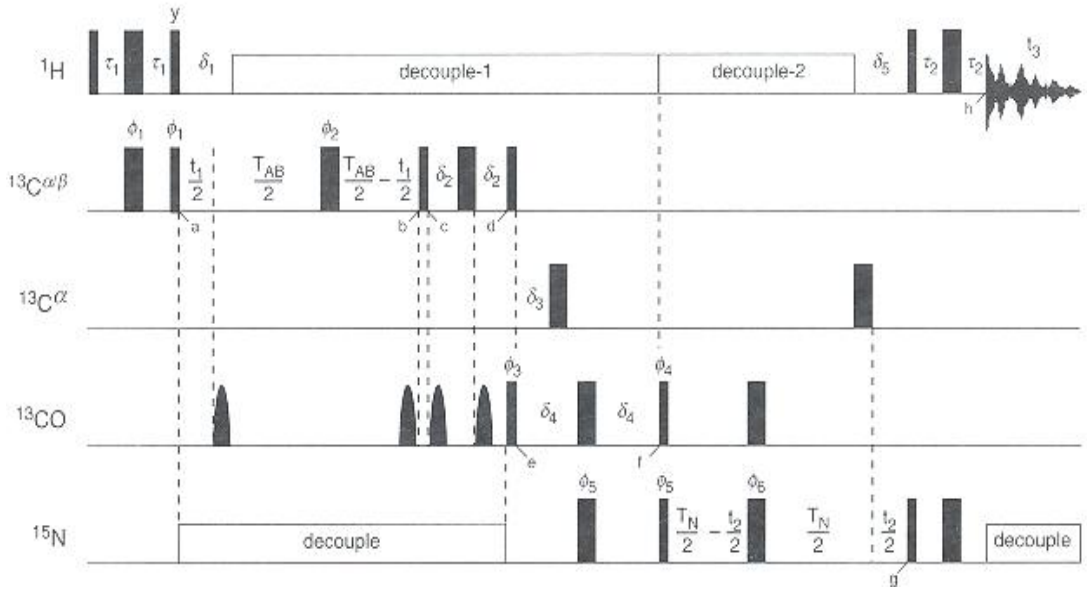


Figure 3.11 Pulse sequence of CBCA(CO)NH. Thin bars represent 90° pulses and thick bars represent 180° pulses. Pulses are applied with *x*-phase unless the phase is indicated above the bar. Rounded bars represent selective 180° pulses applied to ^{13}CO spins[65].

The first INEPT, of duration of 2δ, is used to transfer magnetization from the aliphatic protons H^{α} and H^{β} to the carbons: $H_z^{\alpha} \rightarrow H_z^{\alpha} C_y^{\alpha}$ and $H_z^{\beta} \rightarrow H_z^{\beta} C_y^{\beta}$. It is followed by a constant time evolution period of duration $2T_{AB}$ combined with refocusing of transverse carbon coherence that is antiphase with respect to its attached proton during the delay ε: $H_z^{\alpha,\beta} C_y^{\alpha,\beta} \rightarrow C_x^{\alpha,\beta}$. The time ε should be chosen to be about $1/3J_{HC}$ in order to allow refocusing of CHn for n=1,2,3. Starting from this point decoupling of ^{1}H is necessary. During the $2T_{AB}$ constant time period, the carbon spins are frequency labeled, coupling to carbonyls is refocused but J_{CC} couplings evolve, and the 90° pulse on carbon mixes C^{α} and C^{β} coherences: $C_x^{\beta} \rightarrow C_z^{\beta} C_y^{\alpha}$. The third INEPT of duration 2ζ mixes the aliphatic carbon magnetization with the carbonyl magnetization ($C_z^{\alpha} C'_y$ term is created) and the antiphase coherence $C_y^{\alpha} C_z^{\beta}$ is refocused. In the fourth INEPT, during the period η, the antiphase carbonyl magnetization with respect to C^{α} is refocused ($C^{\alpha}_z C'_y \rightarrow C'_x$) and subsequently transferred to ^{15}N: ($C_z^{\alpha} C'_y \rightarrow C'_z N_y$). The carbonyl and ^{15}N coupling evolves during the period 2θ.

The fifth step consists of constant time evolution ^{15}N during t_2 of duration $2T_N$. The coupling to C^{α} is refocused with 180° pulses on ^{15}N and C^{α}. The ^{15}N magnetization antiphase with respect to carbonyl is refocused ($C'_z N_y \rightarrow N_x$). At this point ^{1}H and ^{15}N decoupling is switched off to allow the nitrogen to become antiphase with respect to ^{1}H:

$N_x \rightarrow N_z H_y$. Finally the last INEPT brings the magnetization to the amide protons which are detected during t_3: $N_z H_y \rightarrow H_x$.

The overall flow of magnetization during the CBCA(CO)NH pulse sequence is schematized as follows:

$$H_{\alpha,\beta} \xrightarrow{{}^1J_{CH}} C_{\alpha,\beta}(t1) \longrightarrow C_{\alpha} \xrightarrow{{}^1J_{NC'}} C' \xrightarrow{{}^1J_{C'N}} {}^{15}NH\ (t2) \xrightarrow{J_{HN}} HN\ (t3)$$

3.3.3 HNCACB

This experiment, together with the CBCA(CO)NH is very important to get a complete sequential assignment of the backbone resonances. Differently from the CBCA(CO)NH, this experiment generates four cross peaks correlating the ^{15}N amide nitrogen, and the H^N amide proton to the C^{α} and C^{β} frequencies of the residues i and i-1.

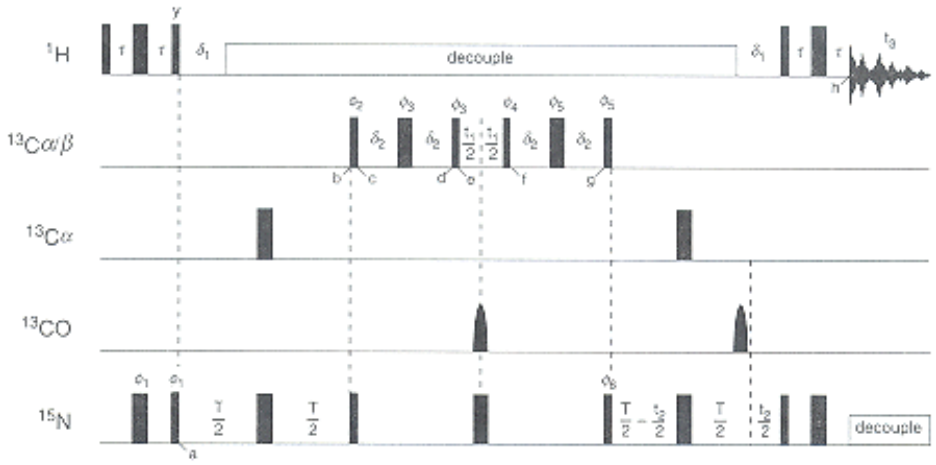
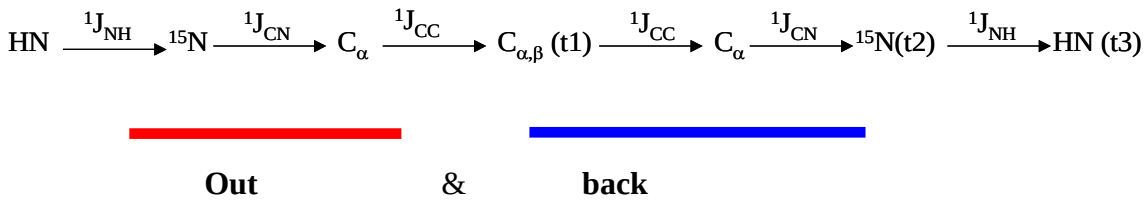


Figure 3.12 Pulse sequence of HNCACB**.** Thin bars represent 90° pulses and thick bars represent 180° pulses. Pulses are applied with *x*-phase unless the phase is indicated above the bar. Rounded bars represent selective 180° pulses applied to ^{13}CO spins[65].

An inspection of the HNCACB pulse sequence (Figure 3.12) reveals an "out and back" mechanism for magnetization transfer between the nuclei. As in previous experiments, the magnetization starts on H^N and is transferred to the other nuclei by INEPT blocks and frequency labeling is performed in constant time units. The pathway of the magnetization transfer:



One disadvantage of this experiment is that carbon chemical shifts recorded during t_1 are modulated by CC-coupling[76].

4. NMR to study protein dynamics and interactions

It is becoming increasingly clear that a detailed understanding of molecular function requires not only high-resolution structural information of the sort that can be obtained through X-ray diffraction or NMR studies but also a quantitative characterization of how structure changes with time. In the case of proteins, in particular, dynamics have been implicated in many aspects of function, including molecular recognition, signaling[77], folding[78], [79], enzymatic catalysis[80], [81] and allostery[82].

In this research project the attention is mainly focused on dynamic processes related to protein-ligand interactions, that could be tightly linked to the BABP biological function concerning the regulation of bile acid concentration within cells and delivery to target receptors.

Proteins are dynamic rather than static systems, and they must perform motions to execute their functions. Motions are possible only if a given protein can assume a large number of somewhat different conformation.

Molecular motions can be exceedingly complex. To accurately describe their time scales, amplitudes and directionalities a very large number of parameters would be required, among these:

1. The line widths dependence on MW and conformational states
2. The proton amide exchange with solvent
3. The number of signals per atom
4. The rates of heteronuclear relaxation processes: (i) R_1 ($1/T_1$) for spin-lattice relaxation; (ii) R_2 ($1/T_2$) for spin-spin relaxation; (iii) Nuclear Overhauser Effect (NOE) for the dipole-dipole relaxation.

Through the measurements of all these parameters, NMR experiments allow the study of conformational motions in a very broad time scale, from picoseconds to seconds. T_1, T_2 and NOE values give information on motions in the range of ps-ms time scale. Relaxation dispersion CPMG data reveal motion in the range of μs-ms time scale. Lineshapes together with chemical shifts analysis detect the slowest motions, in the range of μs to ms for lineshapes and in the range of seconds for chemical shift.

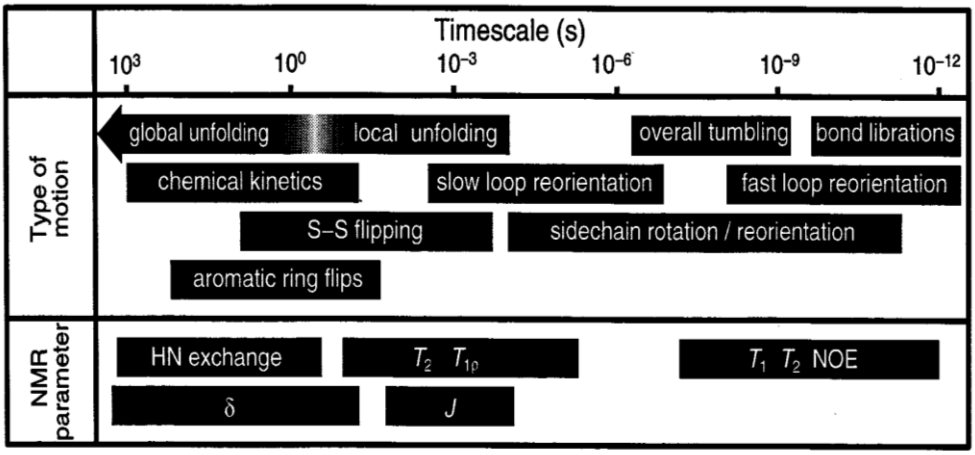


Figure 4.1 Scheme of NMR parameters and motion time scales**.**

All of the mentioned parameters are typically measured using two-dimensional (HSQC type) experiments in which the intensities of peaks are modulated as a function of a time delay placed at a point in the pulse sequence when the relevant relaxation process is active. Each observable relaxation process involves transitions between quantized magnetic energy levels. Such transitions are stimulated by magnetic fields that oscillate at the transition frequencies. Thus, the relaxation rates are determined by the likelihood that the relevant nuclei experience appropriate oscillating magnetic fields. Such oscillating fields in proteins result from the movements of magnetic nuclei relative to each other or relative to the overall permanent field of the NMR magnet. Consequently, relaxation is sensitive to molecular motion.

In any NMR experiment, the bulk magnetization of the sample is perturbed from its equilibrium state (along the *z*-axis, parallel to the permanent magnetic field), and the emitted signal is observed as the sample returns to equilibrium. Relaxation is the process by which nonequilibrium magnetization returns to the equilibrium state (or transforms to another nonequilibrium state). The longitudinal or spin-lattice relaxation rate ($R_1 = 1/T_1$) is the rate constant at which the magnetization recover to thermal equilibrium. The transverse or spin-spin relaxation rate ($R_2 = 1/T_2$) is the rate constant for dephasing of magnetization in the *x*-*y* plane. Relaxation rates $1/T_1$ and $1/T_2$ are often very sensitive indicators of binding of a small ligand to a large molecule, as relaxation rates for ^{1}H and ^{13}C nuclei in small ligand molecules are controlled by molecular tumbling rates and binding will affect relaxation properties[83].

The Nuclear Overhauser effect (NOE) comes from a particular relaxation path, known as cross-relaxation, originated by the dipole-dipole interaction between nuclei close together in the space. The NOE is defined as the fractional change in intensity of one NMR line when another resonance is selectively perturbed. The steady-state heteronuclear {^{1}H}-^{15}N NOEs are determined as a ratio of cross-peak intensities in two experiments, with and without pre-saturation of amide ^{1}H nuclear spins, usually referred to as NOE and NONOE experiments, respectively[84]. These experiments provide the possibility to obtain a complete atomic-level resolution picture of protein dynamics in a time range from picoseconds up to nanoseconds, the range where several motions relevant to protein interaction take place.

In this work, ^{15}N relaxation HSQC type experiments, were based on INEPT magnetization transfer in the ^{1}H-^{15}N pair and contain relaxation-specific pulse sequences inserted during the time period when magnetization is in the ^{15}N dimension. For T_1 and T_2 measurements of apo and complexed cL-BABPs, a series of several 2D spectra were recorded at different resonance fields, with various relaxation delays, ranging from very short to the longest one. The delay values were selected so that they were uniformly distributed over this time interval.

Relaxation times were calculated via least squares fitting of peak intensities, using the rate analysis routine of NMRView program. The heteronuclear NOE effects were calculated from the ratio of cross-peak intensities in spectra collected with and without amide proton saturation. Through the calculation of T_1/T_2 ratio and by the analysis of the NOEs values, it is possible to deduce information about conformational motion of individual residues of the protein. Indeed, high values of T_1/T_2 ratio (higher than the average value plus a standard deviation) are indicative of the presence of slow conformational motions (μs-ms) in NMR time scale, while low values of NOEs (lower than the average value minus a standard deviation) are indicative of the presence of fast conformational motions (ps-ns).

Biological processes, occurring on the μs to ms time scale, depend on molecular dynamics that leads to excursions from highly populated ground states to much less populated excited states. The low populations and the transient formation of excited states are usually invisible to the conventional methods of structural biology and can be studied through NMR relaxation dispersion measurements. The unique capabilities of NMR relaxation dispersion to determine exchange rates on the slow time scale, follow

from the fact that large numbers of residues can be probed individually in a single experiment.

NMR lineshape analysis has been singled out as the method of choice for ligand binding investigation. Indeed NMR lineshape analysis covers a time frame that is highly relevant for the dynamic processes associated to ligand interaction and can be used to determine kinetic rates involving different steps, intermediates or parallel, so to derive kinetic mechanisms. In the case of ligand binding the kinetic analysis is facilitated by the fact that a series of lineshapes can be obtained by varying the concentration of the ligand. The obtained lineshapes, derived from two-dimensional spectra, provide kinetic information on individual residues of the protein.

4.1 Relaxation Dispersion (CPMG)

Here a short description is provided of an NMR technique termed Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion NMR that can give atomic resolution information on millisecond time scale conformational transition occurring in proteins[85]. In cases where conformational exchange occurs between a ground state and “excited” states that are populated at levels of 0.5% or higher with rates of exchange on the order of a few hundred to several thousand per second, the relaxation data measured for multiple nuclei often allow complete kinetic and thermodynamic characterization of the exchange process, even though the resonances of the “excited” states cannot be directly observed in NMR spectra. Additionally, relaxation dispersion data report NMR chemical shift differences between states, providing structural information on the low-populated “excited” species.

4.1.1 CPMG NMR Relaxation Dispersion theory

As clearly described by Kay relaxation dispersion is the dependence of the rate of magnetization decay, $R_{2,\mathrm{eff}}$, during the CPMG sequence, on νCPMG. In practice, $R_{2,\mathrm{eff}}$ rates are quantified from peak intensities in NMR spectra recorded with different numbers of repeats of the τ-180x°-τ CPMG block[86].

Although a complete description of many NMR experiments requires quantum mechanics, much insight can be obtained from a classical picture in which the macroscopic nuclear magnetization at equilibrium is represented by a vector that aligns along an axis (*Z*) that is collinear with the magnetic field, $\mathbf{H}_0$. Figure 4.2a[87]. Application of a 90*y*° radio frequency pulse at the appropriate frequency rotates the magnetization vector along the *X*-axis. Subsequently, magnetization rotates about the *Z*-axis with a characteristic angular frequency Ω, in the example shown in Figure 4.2a, there are two such components (red and blue), each with discrete resonance frequencies. Over time each magnetization component will acquire a phase, φ, with respect to the *X*-axis, given by $\varphi = \Omega t$. In the basic spin–echo experiment (Figure 4.2b) a τ-180*x*°-τ block is added after the 90*y*° pulse, followed by signal acquisition. During the first τ element, magnetization in each of the red and blue states acquires phase $\varphi_A = \Omega_A \tau$ and $\varphi_B = \Omega_B \tau$, respectively, so that a plot of φ vs t is linear, with slope Ω (Figure 4.2b, middle). At $t = \tau$, a "refocusing" 180*x*° pulse is applied that inverts the phase of the magnetization. The phase then "builds-up" again during the second τ period so that the net phase is zero at $t = 2\tau$, giving rise to an ordering of magnetization along the *X*-axis referred to as a spin-echo. Now suppose that the red and blue magnetization components derive from a single NMR active probe (a ^{1}H, ^{13}C, or ^{15}N nucleus, for example) on a molecule that exchanges stochastically between two distinct conformations. If the exchange of conformations leads to a difference in the magnetic environment of the probe, corresponding to resonance frequencies ΩA, ΩB, then a plot of φ vs t might look like one of those shown in Figure 4.2b (bottom), where each magnetization component has acquired a different phase before and after the refocusing pulse, $\varphi(2\tau) \neq 0$. Different $\varphi(t)$ profiles would be observed for different molecules since the process of exchange is random, and because the net signal is given by the sum of $\varphi(2\tau)$ over all molecules, the intensity would be reduced relative to the case of no exchange. CPMG-type dispersion experiments exploit multiple-echo refocusing pulse sequences. In a typical experiment a 90*y*° pulse is followed by an even number of repeats of the τ-180*x*°-τ block (Figure 4.2c). The frequency of application of 180*x*° pulses in the CPMG sequence, $1/(4\tau)$, is called the CPMG frequency, ν_{CPMG}.

Qualitatively, if the rate of molecular exchange is slow compared with ν_{CPMG}, magnetization is nearly completely restored along the *X*-axis after each block, so the amplitude of consecutive spin-echoes is little affected by the exchange process. In

contrast, as the exchange events become on the order of ν_{CPMG}, the CPMG sequence becomes less effective at refocusing, thus leading to a decay of the spin-echo amplitude. NMR relaxation dispersion profiles $R_{2,eff}(\nu CPMG)$ provide sensitive measures of conformational/chemical exchange processes allowing extraction of populations of exchanging states (p_n), rates of transitions between states (k_{mn}), and absolute values of frequency differences between states $|\Delta\omega_{mn}| = |\omega_m - \omega_n|$ (or, equivalently, chemical shift differences in ppm $|\Delta\tilde{\omega}_{mn}|$) for each pair of exchanging states m and n. All of the observables mentioned above can be obtained by a least-squares fit of experimental relaxation dispersion data (i) to approximate expressions (which are available for two-state exchange only) or (ii) by solving the Bloch-McConnell equations numerically for a given chemical exchange model[85]. In most cases, exchange contributions to $R_{2,eff}$ decrease with increasing νCPMG (see Figure 4.2). In the absence of exchange or in the case of $\Delta\tilde{\omega}_{mn} = 0$, $R_{2,eff}$ rates are independent of νCPMG, so that the measured relaxation dispersion profiles are flat.

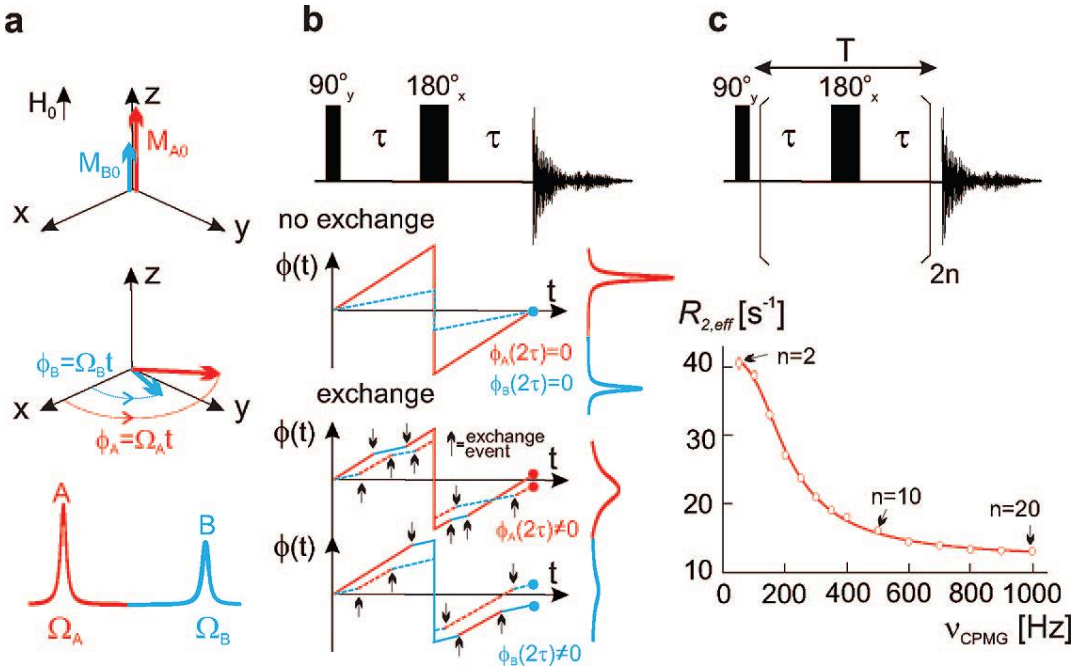


Figure 4.2 The physical basis underlying CPMG relaxation dispersion NMR spectroscopy[87].

4.1.2. Overview of Relaxation Dispersion experiments

In my thesis, relaxation dispersion experiments were recorded at two magnetic fields, on ^{15}N cL-BABP/S-S at different level of saturation by GCDA.

The necessity to acquire these experiments at two magnetic fields, as reported by Millet et al.[88], depends on the nature of the chemical or conformational kinetic processes (on μs-ms). Such processes, generally referred to as chemical exchange contributes to the transverse relaxation rate (R_2), stochastically transfer nuclear spins between magnetic environments with different isotropic chemical shifts.

Our relaxation dispersion data were investigated taking into account a two-site exchange reaction (A ⇌ B), with site A more highly populated than site B ($p_a > p_b$), a shift difference between sites equal to $\Delta\omega$, and an exchange rate constant k_{ex}. The transverse relaxation rate constants for the two sites are given by R_A and R_B, respectively. The exchange contribution to the transverse relaxation rate constant for the more highly populated site is denoted R_{ex} and could be calculated assuming $R_A = R_B$. The timescale of an exchange process can be estimated by comparing R_{ex} at different static (B_0) magnetic fields. This timescale is characterized by a scaling parameter α given by:

$$\alpha = \frac{(B_0 2 + B_0 1)(R_{ex} 2 - R_{ex} 1)}{(B_0 2 - B_0 1)(R_{ex} 2 + R_{ex} 1)}$$

where 1 and 2 are referred to two different static magnetic fields and the value of α depends on the NMR chemical shift time scale for the exchange process:

- for slow exchange ($k_{ex}/\Delta\omega < 1$), $0 \leq \alpha < 1$;
- for intermediate exchange ($k_{ex}/\Delta\omega = 1$), $\alpha = 1$;
- for fast exchange ($k_{ex}/\Delta\omega > 1$), $1 < \alpha \leq 2$.

The dependence of α on a chemical shift time scale is a generalization of the result that chemical exchange broadening is independent of static magnetic field in the slow exchange limit and scales quadratically with the static magnetic field in the fast exchange limit.

Consequently, the static magnetic field dependence of R_{ex} defines the chemical shift time scale for an exchange process even if the populations are so highly skewed

($p_a \gg p_b$) that the minor resonance is not observable in the slow exchange limit considering a two-site exchange reaction.

The variation of $R_{2,\text{eff}}$ as a function of $1/\tau_{cp}$ is called relaxation dispersion. $1/\tau_{cp}$ is the delay between 180° pulses in the CPMG pulse train.

In this work, a fit of $R_{2,\text{eff}}$ values was performed using a master routine in the NMRKIN software, taking into account the possibility to have exchange processes at different timescales. For exchange processes in the range of intermediate to fast time scales, $R_{2,\text{eff}}$ was obtained using the equations reported by Millet et al.[88] (see Appendix E). For exchange processes in slow time scale was used the equation reported by Tollinger et al.[89] (See Appendix E). The values obtained were simultaneously fitted for the two magnetic fields[88], using the two site exchange equation. An example of relaxation dispersion profile measured at two magnetic fields and fit with the general two-site equation is reported in Figure 4.3.

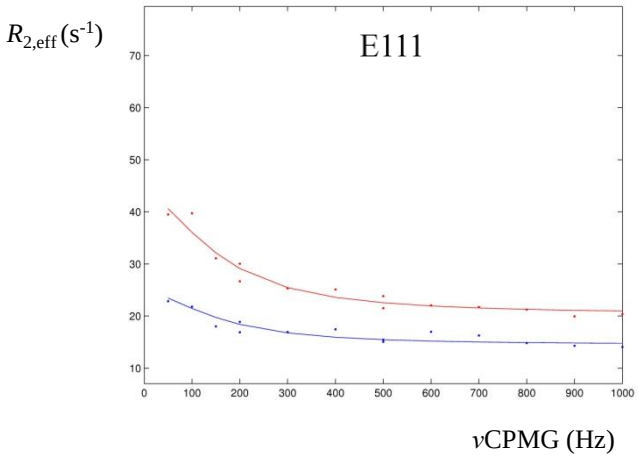


Figure 4.3 Relaxation dispersion profile of residue E111 of apo cL-BABP/S-S, measured at 600 and 900 MHz and fit with the general two-site equation.

The pulse sequence, used in the present work, to record relaxation dispersion spectra, is characterized by constant-time CPMG intervals[86]. In addition, the use of constant-time CPMG periods with variable pulse spacing allowed measurement of a complete dispersion curve in a short measuring time. The CPMG pulse sequence used for ^{15}N cL-BABP/S-S is represented in Figure 4.4.

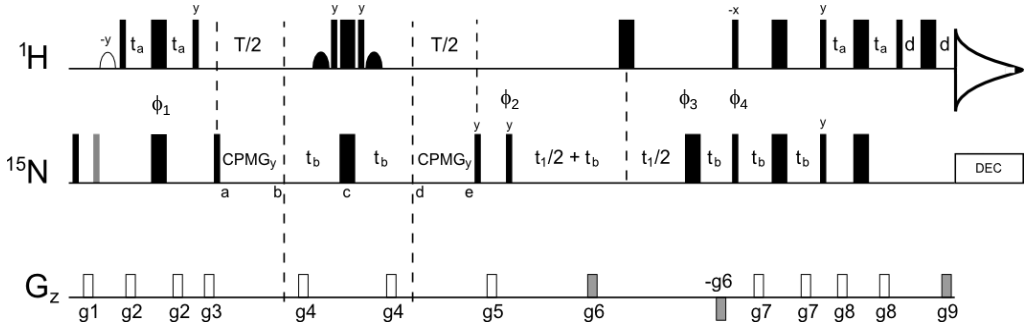


Figure 4.4 Pulse scheme used to measure relaxation dispersion profiles of NH groups in ^{15}N-cLBABP/S-S. All narrow (wide) pulses are applied with a flip angle of 90° (180°) along the *x*-axis, unless indicated otherwise. The shaped ^{1}H pulse at the beginning of the sequence is a 7 ms water selective pulse with the EBURP-1 profile[90]. All rectangular ^{1}H pulses are applied with a field of 30 kHz and are centered on water, following WATERGATE strategy.

CPMG ^{15}N pulses are applied with 6.2 kHz field and the other ^{15}N pulses are applied with 6.8 kHz field. All the ^{15}N pulses are centered around 120 ppm. ^{15}N decoupling is achieved using a WALTZ-16 field[91] (1.2 kHz field at both 600 and 900 MHz).

Each CPMG*y* element is of the form (τCPMG-180°-τCPMG)*n*/2, where $T = n\,(2\tau\text{CPMG} + pw_{N180})$, *n*/2 is even, and pw_{N180} is the ^{15}N 180° pulse width. Quadrature detection in F_1 is achieved using the enhanced sensitivity approach[92], [93] where for each value of t_1 separate data sets are recorded for (ϕ4,g9) and (ϕ4+180°,-g9). The delays used are as follows: $\tau a < 1/4J = 2.25$ ms; $\tau b = 1/(4J) = 2.68$ ms, $T = 40$ ms (νCPMG =1000 Hz). The phase cycle is: $\phi 1 = (x,-x)$; $\phi 2 = 2(x),2(-x)$; $\phi 3 = 2(x),2(-x)$; $\phi 4 = x$; rec = $x,2(-x),x$. For each successive t_1 point $\phi 2$ is incremented by 180° in concert with the receiver[94]. The durations and strengths of the gradients are: $g1$ = (1.0 ms, 5 G/cm); $g2$ = (0.5 ms, 4 G/cm); $g3$ = (1.0 ms,-6 G/cm); $g4$ = (0.5 ms, 18 G/cm); $g5$ = (0.8 ms, 15 G/cm); $g6$ = (1.1 ms, 15 G/cm); $g7$ = (0.4 ms, 3.5 G/cm); $g8$ = (0.3ms, 2.5 G/cm); $g9$ = (0.11 ms, 29.4 G/cm)[86]. Gradient coherence selection achieved with shaded gradients.

During the interval extending from *a* to *e* in the sequence a fixed number of ^{15}N refocusing pulses were applied during each of the two *constant time* CPMG intervals of equal length (*a* to *b* and *d* to *e*). A series of 2D spectra are recorded as a function of effective rf field strength by varying the number of 180° pulses and therefore the pulse spacing so that the net relaxation time (*a* - *b*, *d* - *e*) is the same in each experiment. The intensities of cross-peaks in 2D spectra recorded for a given rf field, νCPMG, are converted into decay rates, $R_{2,\text{eff}}$, via

$$R_{2,\text{eff}}(\nu\text{CPMG}) = \frac{-1}{T}\ln\frac{I(\nu\text{CPMG})}{I_0}$$

Where $I(\nu\text{CPMG})$ and I_0 are the intensities of a given cross-peak with and without the CPMG periods from *a* to *b* and *d* to *e* in Figure 4.4, *T*/2 is the length of each CPMG

train and ($\nu CPMG = 4\tau CPMG$), with $2\tau CPMG$ the separation between the centers of successive refocusing pulses. A plot of $R_{2,eff}$ as a function of field strength, $\nu CPMG$, gives a relaxation dispersion profile for each residue from which exchange parameters can be obtained[86].

4.2 Lineshape analysis in protein interactions

Dynamics processes on the μs to ms time scale influence the shape of the NMR signal if they are active during the acquisition of the free induction decay (FID). NMR lineshapes can therefore be used to derive kinetic mechanisms and to determine kinetic rates

For a protein ligand binding process the relevant factors are the rates of exchange between the unbound, complexed and possible intermediate forms and the populations of the species involved in the exchange process. It is important to note that exchange observed by NMR for a protein ligand interaction does not necessarily represent the rate of ligand binding release but rather the relaxation of the protein from one form into another. Since binding processes are usually diffusion controlled they don't fall into the exchange window which influences NMR lineshapes and it is therefore the release process that is observed by NMR[2].

4.2.1 Mathematical treatment of NMR signals

The effect of exchange processes on NMR signals based on Bloch equations (Bloch, 1946) was first derived by Gutowsky et al (1953) solving the following equation:

$$\mathrm{M} = \frac{i\omega_1 M_0 \left[(\tau_A + \tau_B) + \tau_A\tau_B\left(\alpha_A p_B + \alpha_B p_A\right]\right.}{(1 + \alpha_A\tau_A)(1 + \alpha_B\tau_B) - 1} \quad (1)$$

Here M is the macroscopic magnetization and M_0 is the magnetization at the equilibrium or at the steady-state. With $\alpha_A = T_2^{-1} - i\,(\Delta\omega + \delta\omega/2)$ and $\alpha_B = T_2^{-1} - i\,(\Delta\omega + \delta\omega/2)$, $\tau_A = {}^{1}/_{k_A} = {}^{\tau}/_{p_B}$ and $\tau_B = {}^{1}/_{k_B} = {}^{\tau}/_{p_A}$.

Considering that $\Delta\omega = 2\pi\Delta\nu$, $\Delta\nu$ is the separation of lines in Hz and $\delta\omega/2$ is the width of the signal at half height. The NMR signal experiences the most extreme broadening signal at the coalescence point where exchange rate is $k_{coal} = \Delta\omega$;

For $k >> \Delta\omega$, on a fast exchange on NMR timescale, a single signal is observed at an intermediate chemical shift between that of the two exchanging nuclei;

For $k << \Delta\omega$, on a slow exchange on NMR timescale, two signals are visible for the exchanging nuclei (see Figure 4.4).

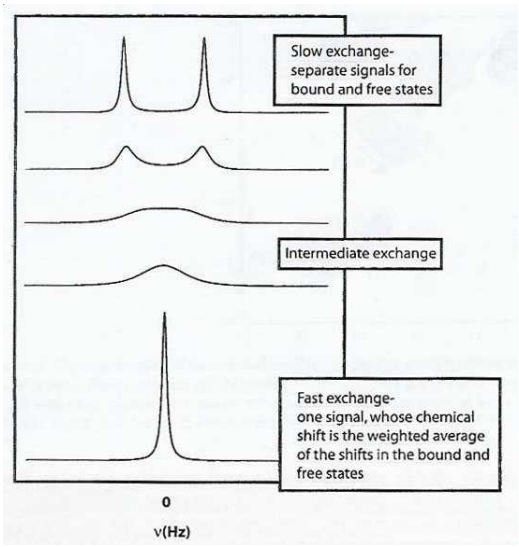


Figure 4.4 Effects of the exchange rate on the appearance of the NMR spectrum from a system in which a nucleus is exchanging between two equally populated states in which it has a different chemical shift[95].

For the exchange broadening of lines $LW_{\mathrm{A}} = LW_{\mathrm{obs}} - LW_{\mathrm{nat}}$ (where the natural line width $LW_{\mathrm{nat}} = \frac{1}{\pi T_2^*}$ includes inhomogeneity broadening) the following approximations are commonly used:

slow exchange: $1/\tau_A = k_A = \pi\, LW_{\mathrm{A}}$

fast exchange: $1/\tau_A = k_A = \frac{\Delta\omega^2 p_A p_B^2}{LW_{\mathrm{A}}}$

To study exchange between multiple sites or reactions involving more complex kinetic mechanisms, the Bloch equations, modified for chemical exchange, under steady state conditions, can be represented in a matrix notation (Binsch, 1968) that has been included in the NMRKIN software[2] as:

$$\mathbf{AM} = i\mathrm{C}\mathbf{P} \quad (2)$$

$\boldsymbol{M}$ is a column vector containing the magnetizations M_i of the nuclei which are subject to chemical exchange;

P is a column vector with populations (mole fractions) of the reaction components;

$C = B_1 M_0$;

A is a square matrix that contains the:

$$\mathbf{A} = 2\pi i\ (\mathbf{I}\nu - \mathbf{W}_0) + \mathbf{R}_2 + \mathbf{K} \quad (3)$$

I is the unit matrix;

ν is a variable frequency defining the frequency range;

$\mathbf{W}_0$ is a diagonal matrix with Larmor frequencies ν_i;

$\mathbf{R}_2$ is a diagonal matrix with transverse relaxation rates, $R = 1/T_2^* = \pi$ LW (line width at half height);

K is a matrix containing the exchange rates.

The total magnetization is the sum of the individual magnetization components:

$$M_{tot} = \mathbf{I}^T\mathbf{M} = iC1^T\mathbf{A}^{-1}\mathbf{P} \quad (4)$$

To calculate lineshape it is necessary the inversion of the matrix **A** for each point of the spectrum. From equation (4) the corresponding frequency is calculated:

$$F(t) = \mathbf{I}\exp\ [(-\mathbf{K} + 2\pi i\mathbf{W}_0 - \mathbf{R}_2)\ t]\ \mathbf{P} \quad (5)$$

After Fourier transformation the obtained spectra is the same obtained for equation (4). With this procedure there is the advantage that typical processing steps as zero filling or apodization functions can be incorporated into lineshape simulation.

4.2.2 Lineshape simulation through NMRKIN

The matrix notation used till now was implemented within MATLAB language to create the routines of NMRKIN[2] necessary for lineshapes cross-sections simulations. Input data for NMRKIN software should be a MATLAB matrix or, as in this case, an ASCII format directly generated from NMRlab, an NMR processing software implemented in MATLAB[59].

NMRKIN uses the simulation of the proton and the nitrogen dimensions of HSQC spectra that allows calculation of off-rates for protein-ligand interaction observed at individual amino acids of the protein. The simultaneous simulation of the cross-sections of both dimension give the possibility to consider the effect of the line broadening when it occurs in one of the two dimension.

When the program is started a setup file is requested to determine all important parameters.

Required input parameters:

LW: line width

w: position of the line

τ_{21} "off-lifetime"

τ_{23} "off-lifetime"

p_A and p_B or p_C because $p_{TOT} = 1$

In the simulation identical exchange parameters (**P**, and k_{off}) and different peak separations (Δ_{vN} and Δ_{vH}) are considered. The relative intensity of lines in the second dimension are used to scale the intensity of lines in the first dimension and vice versa.

Parameter adjustment can be performed manually, as in this work, or by a least squares routine available in the MATLAB optimization toolbox.

NMRKIN software used for lineshape simulations includes a series of mechanism that are commonly observed in proteins and allows the implementation of equations corresponding to additional mechanisms in the form of MATLAB scripts. The most common mechanism of protein ligand interaction can be viewed as bimolecular second order interaction between a protein and a ligand. More complex mechanisms could involve intermediates or parallel binding reactions in which different conformers of a protein, existing in parallel in the free form, converge to a single end point in the

protein-ligand complex. Many more complex mechanisms are theoretically feasible although it is usually very difficult to fit calculated lineshapes to the experimental ones.

4.2.3 One step binding mechanism

One of the mechanisms described by NMRKIN is the bimolecular simple one step second order reaction that could be represented by the equation:

$$\mathrm{P} + \mathrm{L_n} \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} \mathrm{PL_1L_2}$$

employing a two-state binding mechanism.

The exchange matrix used in NMRKIN to describe the reported binding mechanism is :

$$\mathbf{K} = \begin{bmatrix} k_{12}\,[\mathrm{L}] & -k_{21} \\ -k_{21} & k_{12}\,[\mathrm{L}] \end{bmatrix} = \begin{bmatrix} k_{12}^{'} & -k_{21} \\ -k_{21} & k_{12}^{'} \end{bmatrix}$$

Exchange between two states is described by mole fractions (= the populations p_i) of one of the two states ($p(P) = 1 - p(\mathrm{PL_1L_2})$), the line widths of the two signals and their frequencies separations ($\Delta\nu_{P-\mathrm{PL_1L_2}}$).

Here the position of the equilibrium between the two states depends on the affinity of the ligand L for the two protein binding sites and on amount of L that has been added. At the equilibrium $\frac{d[\mathrm{P}]}{dt} = \frac{d\,[PL_1L_2]}{dt} = 0$; and therefore $k_{12}[\mathrm{L}]p_{\mathrm{P}} = k_{21}p_{\mathrm{PL_1L_2}}$. The effective on-rate $k_{12}^{'} = k_{12}\,[\mathrm{L}]$ depends on the concentration of the ligand L, unknown at the equilibrium.

This bimolecular mechanism is referred in Chapter 8 as m21 binding mechanism.

4.2.4 Two step binding mechanism

In this work, a mechanism for two-stages binding was implemented to simulate the lineshapes of cL-BABPs interacting with two bile acid molecules.

The kinetic model was written taking into account (i) the presence of a generic intermediate that could arise from different types of bound species; (ii) the presence of two binding sites, (iii) the dependence of both binding steps on the ligand concentration and (iv) the independence of the first on-rate from the second one and vice versa.

The implemented mechanism is described by

$$\mathrm{P} + \mathrm{L} \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} \mathrm{PL_n} + \mathrm{L} \underset{k_{32}}{\overset{k_{23}}{\rightleftarrows}} \mathrm{PL_1L_2}$$

In this model the on-rate and the off-rate of the two binding steps are different and it is assumed that second ligand binds sequentially after the formation of the intermediate PL.

Furthermore it is assumed that there is no direct conversion of the free protein P into $\mathrm{PL_1L_2}$, i.e. the rates k_{13} and k_{31} have a value of zero.

The rate matrix for this case is:

$$\mathbf{K} = \begin{bmatrix} k_{12}^{'} + k_{13} & -k_{21} & -k_{31} \\ -k_{12}^{'} & k_{21} + k_{23}^{'} & -k_{32} \\ -k_{13} & -k_{23}^{'} & k_{31} + k_{32} \end{bmatrix} \equiv \begin{bmatrix} k_{12}^{'} & -k_{21} & 0 \\ -k_{12}^{'} & k_{21} + k_{23}^{'} & -k_{32} \\ 0 & -k_{23}^{'} & k_{32} \end{bmatrix}$$

Here $k_{12}^{'} = k_{12}[\mathrm{L}] = \frac{1}{\tau_{12}}$ and $k_{23}^{'} = k_{23}[\mathrm{L_1L_2}] = \frac{1}{\tau_{23}}$ depend on the concentration of the added ligand.

The matrix can be also written as

$$\mathbf{K} = \begin{bmatrix} 1/\tau_{12} & -1/\tau_{21} & 0 \\ -1/\tau_{12} & 1/\tau_{21} + 1/\tau_{23} & -1/\tau_{32} \\ 0 & -1/\tau_{23} & 1/\tau_{32} \end{bmatrix}$$

Here $\frac{1}{\tau_{21}} = k_{21} = k_{\mathrm{off},1}$ and $\frac{1}{\tau_{32}} = k_{32} = k_{\mathrm{off},2}$.

For equilibrium conditions we get:

$$k_{12}[\mathrm{L}]p(\mathrm{P}) = k_{21}p(\mathrm{PL_n})$$

$$\tau_{21}p(\mathrm{P}) = \tau_{12}p(\mathrm{PL_n})$$

and therefore

$$\tau_{12} = \tau_{21}\frac{p(\mathrm{P})}{p(\mathrm{PL_n})} \quad (1)$$

And the same for the second step of the mechanism:

$$k_{23}[\mathrm{L}]p(\mathrm{PL_n}) = k_{32}p(\mathrm{PL_1L_2})$$

$$\tau_{32}p(\mathrm{PL_n}) = \tau_{23}p(\mathrm{PL_1L_2})$$

and therefore

$$\tau_{23} = \tau_{32}\frac{p(\mathrm{PL_n})}{p(\mathrm{PL_1L_2})} \quad (2)$$

After eliminating $p(\mathrm{PL_n})$ from equation (1) and (2) we get

$$\tau_{21}\tau_{32}\frac{p(\mathrm{P})}{p(\mathrm{PL_1L_2})} = \tau_{23}\tau_{12}$$

With $p(P) + p(\mathrm{PL_n}) + p(\mathrm{PL_1L_2}) = 1$

$$\tau_{23} = \tau_{32}\frac{1 - p(\mathrm{PL_1L_2}) - p(\mathrm{P})}{p(\mathrm{PL_1L_2})}$$

The NMRKIN mechanism file was set up with these simplifications and is referred in Chapter 8 as m2121 binding mechanism.

5. Apo and holo cL-BABP/S-S resonance assignment

The assignment of the NMR resonances of the cL-BABP/S-S apo and in complex with different bile acids, is the first step to study the effect of the disulphide bond introduction on the protein structure.

The cL-BABP/S-S samples used for the assignment were enriched in ^{15}N or ^{15}N and ^{13}C. The protein-ligand complexes were in a ratio 1:3 for the homotypic complexes with unlabeled CDA or GCDA and in a ratio 1:2:2 for the heterotypic complex with the unlabeled GCA-GCDA (see Tables 1-4).

In ^{1}H-^{15}N HSQC spectra every cross-peak corresponds to an amide group and thereby correlates an amide nitrogen with its covalently bound proton (Figure 5.1).

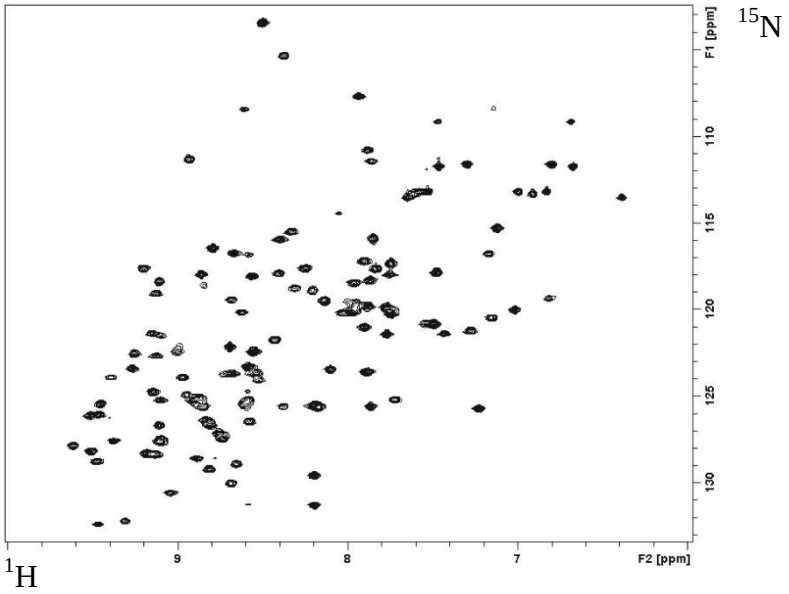


Figure 5.1 ^{1}H-^{15}N HSQC spectrum of the cL-BABP/S-S recorded at 500 MHz at 298K. Each cross-peak represents an amide proton. The correlated proton and nitrogen shifts (frequencies) can be extrapolated from the two axes, where F1 is the first dimension (the amide proton dimension) and F2 is the second, indirectly detected dimension (the nitrogen dimension).

The backbone resonance assignment of signals was performed, whenever possible, comparing the 2D ^{1}H-^{15}N HSQC spectra of cL-BABP with and without disulphide bond. ^{1}H, ^{15}N chemical shifts changes observed in cL-BABP/S-S spectra, with respect to the ones of cL-BABP, gave a first information about the effect of disulphide bond introduction.

In this way only 36% of the backbone resonances of cL-BABP/S-S apo and 77% of the cL-BABP/S-S complexed with CDA were assigned. To implement the assignment a combination of 2D and 3D TOCSY and NOESY HSQC recorded at pH 5.6 and pH 7.2 for the apo form and at pH 7.2 for the various holo forms were acquired, thus obtaining the assignment of ~90% of the ^{1}H, ^{15}N resonances.

The assignment was finally completed using 3D heteronuclear triple resonances NMR experiments: CBCA(CO)NH/HNCACB and HNCO, recorded on the ^{15}N, ^{13}C-labelled protein samples. CBCA(CO)NH/HNCACB experiments allowed to extend the correlations among the amide nitrogen and its attached proton, to Cα and Cβ atoms through bonds. In these 3D heteronuclear experiments the 3rd dimension corresponded to the carbon frequency while the first two dimensions corresponded to the amide proton (HN) and the amide nitrogen (N) frequencies. The plane described by these two axes depicts the reference ^{1}H-^{15}N-HSQC spectrum. Cα and Cβ signals from the (i) and the (i-1) residue, which are transferred to the (i)-amide proton in HNCACB, were correlated with matching carbon signals of the (i-1) residue, transferred to the (i)-amide proton, observed in CACBCONH. The redundancy of information in 3D experiments allowed overlapping clusters to be placed in sequential order and then matched to specific stretches of amino acid residues within the protein sequence.

ASSIGNMENT PERCENTAGE (%)	
cL-BABP/S-S Apo	93% HN,N - 98% Cα,Cβ
cL-BABP/S-S:CDA (1:3)	93% HN,N - 98% Cα,Cβ
cL-BABP/S-S:GCA:GCDA (1:2:2)	93% HN,N - 98% Cα,Cβ - 98% CO
cL-BABP/S-S:GCDA (1:3)	92% HN,N

Table 5.1 Percentages of resonance assignment of Apo cL-BABP/S-S and cL-BABP/S-S complexed with CDA, GCDA/GCA, GCDA.

The assignment of HN, N, and where detected, of Cα, Cβ and CO, are reported in Tables 1-4.

Protein assignments of apo and holo (CDA) have been deposited in the Biological Magnetic Resonance Bank (BMRB), with accession Number 16310 and 16309.

6. The disulphide bridge: relevance and influence on BABP dynamics and interaction properties

- The effect of the presence of the disulphide bridge (S-S) on ligand binding and backbone dynamics of cL-BABPs has been investigated in the following paper [*FEBS Journal*, 276, 6011-6023, (2009)] by combining different labeling strategies and appropriate NMR experiments on protein complexes with GCDA and GCA.
- The obtained data clearly show that the 1:2 protein:ligand stoichiometry is conserved for both ligands in the presence of disulphide bridge, as confirmed also by mass spectroscopy analysis.
- To study exchange regimes among the bound and unbound ligands, 2D ^{1}H-^{15}N HSQC spectra, at different temperatures, and DOSY experiments were acquired. The exchange between the superficial binding site and the free ligand is the most relevant process, independently by the presence/absence of the disulphide bridge. Instead the presence of the (S-S) bond modulate the affinity of GCA for cL-BABP.
- NMR competition experiments clearly highlighted a site-selectivity for GCDA and GCA, not observable in the absence of the (S-S). Other data indicated that GCDA preferentially populates the superficial site while GCA binds to internal site. GCDA seems to assume a different orientation in homotypic complex or when GCA is bound to the internal site.
- Changes in motion propagation within the β-barrel, induced by the (S-S), have been mapped onto the cL-BABP apo structure and the effect of the binding of GCDA and GCA on backbone conformation and mobility has been assessed. ^{15}N Relaxation experiments and chemical shift perturbation (CSP) analysis were performed for apo cL-BABP/S-S, for cL-BABP/S-S complexed with CDA, GCDA, GCA, and for the heterotypic complex with GCDA-GCA. Protein dynamics is largely influenced by the presence/absence of the disulphide bridge. Indeed the presence of the (S-S) bridge favors the propagation of slow motions from the C-terminal region of the molecule to the N-terminal β-sheet in the apo protein, and enhanced backbone motions in cL-BABP/S-S-CDA complex.

My contribution to the following publication, titled: "Disulphide bridge regulates ligand binding site-selectivity in liver bile acid binding proteins", consisted of:

1. Planning of the NMR experiments for the backbone assignment of cL-BABP/S-S apo and holo forms;
2. Backbone assignment of all the proteins described in the paper;
3. Acquisition and analysis of ^{15}N relaxation experiments;
4. Acquisition and analysis of all the titration experiments;
5. Preparation of introduction, results and material and methods section of the paper and collaboration in writing the discussion paragraph.

the FEBS Journal

Disulfide bridge regulates ligand-binding site selectivity in liver bile acid-binding proteins

Clelia Cogliati[1], Simona Tomaselli[1], Michael Assfalg[2], Massimo Pedò[2], Pasquale Ferranti[3], Lucia Zetta[1], Henriette Molinari[2] and Laura Ragona[1]

1 Laboratorio NMR, Istituto per lo Studio delle Macromolecole, CNR, Milan, Italy
2 Dipartimento di Biotecnologie, Università di Verona Strada le Grazie, Verona, Italy
3 Dipartimento di Scienza degli Alimenti, Università di Napoli Federico II, Portici, Italy

Keywords
backbone dynamics; disulfide bridge; intracellular lipid-binding protein; molecular recognition; NMR

Correspondence
L. Ragona, Lab. NMR, Istituto per lo Studio delle Macromolecole, CNR, Via Bassini, 15, 20133, Milano, Italy
Fax: +39 02 23699620
Tel: +39 02 23699619
E-mail: laura.ragona@ismac.cnr.it
H. Molinari, Dipartimento di Biotecnologie, Università degli Studi di Verona, Strada le Grazie, 15, 37134 Verona, Italy
Fax: +39 0458027929
Tel: +39 0458027901
E-mail: henriette.molinari@univr.it

(Received 3 July 2009, revised 17 August 2009, accepted 18 August 2009)

doi:10.1111/j.1742-4658.2009.07309.x

Bile acid-binding proteins (BABPs) are cytosolic lipid chaperones that play central roles in driving bile flow, as well as in the adaptation to various pathological conditions, contributing to the maintenance of bile acid homeostasis and functional distribution within the cell. Understanding the mode of binding of bile acids with their cytoplasmic transporters is a key issue in providing a model for the mechanism of their transfer from the cytoplasm to the nucleus, for delivery to nuclear receptors. A number of factors have been shown to modulate bile salt selectivity, stoichiometry, and affinity of binding to BABPs, e.g. chemistry of the ligand, protein plasticity and, possibly, the formation of disulfide bridges. Here, the effects of the presence of a naturally occurring disulfide bridge on liver BABP ligand-binding properties and backbone dynamics have been investigated by NMR. Interestingly, the disulfide bridge does not modify the protein-binding stoichiometry, but has a key role in modulating recognition at both sites, inducing site selectivity for glycocholic and glycochenodeoxycholic acid. Protein conformational changes following the introduction of a disulfide bridge are small and located around the inner binding site, whereas significant changes in backbone motions are observed for several residues distributed over the entire protein, both in the apo form and in the holo form. Site selectivity appears, therefore, to be dependent on protein mobility rather than being governed by steric factors. The detected properties further establish a parallelism with the behaviour of human ileal BABP, substantiating the proposal that BABPs have parallel functions in hepatocytes and enterocytes.

Introduction

Bile acids (BAs) are vital components of many biological processes and play an important role in the pathogenesis of numerous common diseases [1], but the specific mechanisms coupling intracellular BAs to biological targets are not well understood. BAs circulate between the liver and intestine through a mechanism known as 'enterohepatic circulation', which is a tightly regulated process, particularly by BAs themselves. BA-binding proteins (BABPs), belonging to the intracellular lipid-binding protein (iLBP) family, play a vital role in the enterohepatic circulation as cytoplasmatic transporters of BAs. Understanding the mecha-

Abbreviations
BA, bile acid; BABP, bile acid-binding protein; CA, cholate; CDA, chenodeoxycholate; CSP, chemical shift perturbation; GCA, glycocholic acid; GCDA, glycochenodeoxycholic acid; I-BABP, human ileal bile acid-binding protein; iLBP, intracellular lipid-binding protein; L-BABP, chicken liver bile acid-binding protein.

nism regulating these interactions is a key step in providing a model for the transfer of BAs from the cytoplasm to the nucleus for delivery to nuclear receptors, and can be used to inspire the design of therapeutic agents for the treatment of metabolic disorders, such as obesity, type 2 diabetes, hyperlipidaemia, and atherosclerosis [1–3].

BABPs are characterized by a conserved β-barrel structure, formed by two orthogonal β-sheets, and a helix–loop–helix motif defining, with flexible loops, the so-called protein open end, delimiting the entrance to the barrel cavity. BABPs from various organisms have been shown to bind bile salts with differences in ligand selectivity, binding affinity, stoichiometry, and binding mechanism. The two most extensively characterized BABPs, namely human ileal BABP (I-BABP) and chicken liver BABP (L-BABP), share the common property of binding two bile salt molecules with weak intrinsic affinities and strong positive cooperativity [4–6]. I-BABP, unlike L-BABP, displays remarkable site selectivity for the two main glycoconjugated BAs, glycocholic acid (GCA) and glycochenodeoxycholic (GCDA). A number of factors have been shown to modulate ligand binding, e.g. the chemistry of the ligand and the nature of the protein residues [7,8]. A prominent role for protein plasticity was suggested for L-BABP, where binding was found to be regulated by a dynamic process and accompanied by a global conformational rearrangement [9]. Essential dynamics analysis of the molecular dynamics trajectories obtained for L-BABP indicated that the portal area is the region mostly affected by complex formation, and that the major concerted motions involve the structural elements of the open end, which are dynamically coupled in different ways, whether in the presence or in the absence of the ligands [10]. Another source of ligand-binding variability may be introduced by the presence of disulfide bridges. Indeed, several cases have been reported in the literature for members of the iLBP family where the introduction/removal of a disulfide bridge was responsible for changes in ligand-binding stoichiometry and affinities. The removal of a disulfide bond in rat lipocalin-type prostaglandin D synthase slightly increased the binding affinity for biological ligands, by leading to a less compact barrel pocket and allowing a higher number of residues to contribute to ligand binding [11]. In the cellular retinoic acid-binding protein I, the introduction of a disulfide bond abolished the structural mobility of the portal region, thus leading to irreversible retinoic acid binding [12].

Most liver BABPs belonging to nonmammalian species have a disulfide bond involving the conserved Cys80 and the cysteine at position 91. For L-BABP, two forms are known, in which residue 91 can be either a threonine or a cysteine, although all the studies presented up to now have dealt with the form devoid of the disulfide bridge [5,9,13–15]. The presence of a disulfide bridge in the protein scaffold of the homologous liver zebrafish BABP (69.8% identity, calculated with CLUSTALW) was correlated with the binding stoichiometry [16], which varied from one ligand molecule, with a disulfide bridge, to two ligand molecules, with the Cys80–Cys91 disulfide bridge removed. On this basis, in a continuous effort to establish the determinants of binding stoichiometry and site selectivity in this protein family, the T91C L-BABP protein, with a Cys80–Cys91 disulfide bridge, has been studied by different NMR and MS approaches.

The role of the disulfide bridge in ligand binding and the backbone dynamics of L-BABP has been investigated here by combining different labelling strategies for both ligand and protein with appropriate NMR experiments. The results clearly show that, although the binding stoichiometry is conserved, site selectivity for GCDA and GCA, which is not observable in the absence of the disulfide bridge, is now present. Changes in motion propagation within the β-barrel, induced by the disulfide bridge, have been mapped onto the BABP apo structure, and the effects of the binding of the two most abundant glycoconjugated bile salts on the backbone conformation and dynamics have been clearly assessed.

Results

Effect of disulfide bridge on binding properties

Binding site occupancies

$^{1}H/^{15}N$-HSQC spectra were collected on isotopically enriched physiological glycine conjugates, GCA and GCDA (differing only in the presence of a hydroxyl group at position 12; Fig. S1), complexed with unlabelled T91C L-BABP at different protein/ligand ratios (1 : 0.3, 1 : 0.6, 1 : 1, 1 : 1.5, 1 : 2, 1 : 2.5, and 1 : 3), in order to monitor the number and occupancy of individual binding sites. The spectra obtained for [^{15}N]GCDA revealed the presence of two main resonances, corresponding to [^{15}N]GCDA bound to two distinct binding sites, denoted site 1 (7.17, 117.3 p.p.m.) and site 2 (6.0, 117.5 p.p.m.), whose chemical shifts did not change during the titration, suggesting the presence of a slow exchange regime (Fig. 1A). A few other cross-peaks with chemical shifts very close to those of peak 1 and peak 2 were visible, and were ascribed to heterogeneous binding at site 1

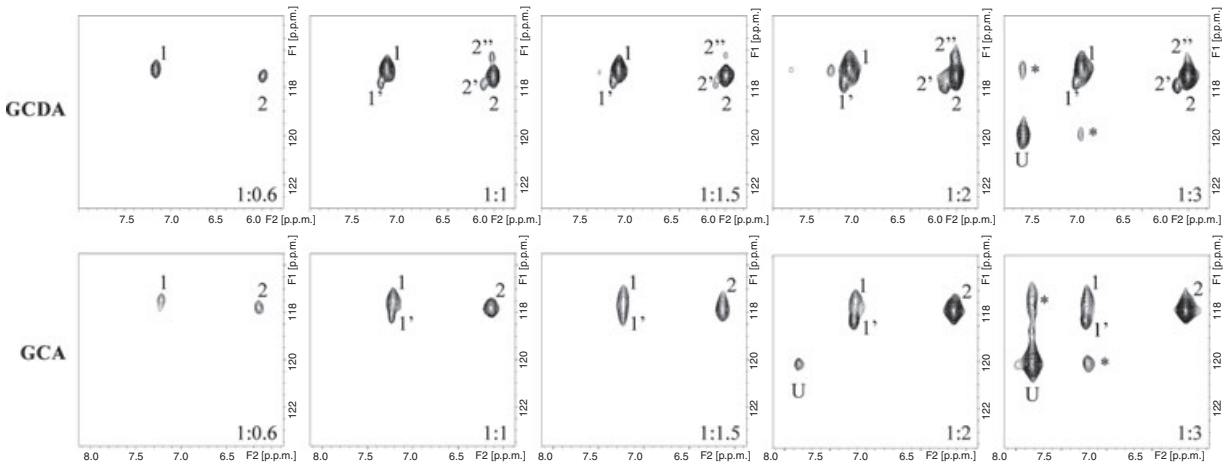


Fig. 1. [^{15}N]GCDA and [^{15}N]GCA in complex with T91C L-BABP. 2D ^{1}H/^{15}N-HSQC spectra at different protein/ligand ratios (1 : 0.6, 1 : 1, 1 : 1.5, 1 : 2, and 1 : 3) were recorded at 298 K at 500 MHz. The resonances corresponding to the unbound ligand and to binding sites 1 and 2 are indicated as U, 1, and 2, respectively. The satellite peaks of site 1 and site 2 are also marked. Asterisks indicate exchange peaks.

and site 2. The unbound resonance (7.8, 119.8 p.p.m.) was visible at protein/ligand ratios higher than 1 : 2, together with exchange peaks between the unbound and site 1 cross-peaks. During the titration, site 1 and site 2 ^{1}H linewidths were substantially unchanged (Fig. 2A). The quantitative volume analysis of these predominant forms indicated that binding site occupancies reached a plateau value, for both sites, at a protein/ligand ratio of 1 : 2 (Fig. 2B). The NMR data thus indicate that, even in the presence of the disulfide bridge, L-BABP maintains the ability to bind two GCDA molecules, at variance with the homologous zebrafish protein. This result was corroborated by MS analysis of T91C L-BABP in complex with GCDA, indicating the presence of the doubly ligated form in solution at a protein/ligand ratio 1 : 2 (data not shown). Similar NMR results were obtained for GCA, and ^{1}H/^{15}N-HSQC NMR titration experiments, performed on the unlabelled T91C L-BABP with increasing amounts of [^{15}N]GCA, indicated the presence of the three cross-peaks named site 1 (7.2, 117.5 p.p.m.), site 1′ (7.2 and 118.0 p.p.m.), and site 2 (6.122, 117.81 p.p.m.) (Fig. 1B). The cross-peak annotated as site 1′ was probably due to the presence of slightly different populations of GCA at this site. The resonance corresponding to the unbound ligand became visible at a protein/ligand ratio of 1 : 2 (7.8 and 120.1 p.p.m.) and exhibited exchange cross-peaks with site 1. The chemical shifts of GCA resonances did not change during the titration, whereas for some of them a variation in linewidth was observed (Fig. 2C), suggesting the presence of a slow to intermediate exchange regime. Site 2 and free GCA resonances exhibited a linewidth decrease upon an increase in protein/ligand ratio. This behaviour is consistent with exchange with free ligand being abolished as saturation is approached [17]. The changes in linewidths did not allow a quantitative determination of site 2 occupancy. The site 1 linewidth ($\sim$ 33 Hz), which was broader than that of site 1′ ($\sim$ 22 Hz), is attributable to exchange with free ligand, as supported by the observation of exchange peaks for site 1 and unbound GCA. Both site 1 and site 1′ linewidths did not decrease as saturation was approached, thus confirming the presence of conformational heterogeneities of the bound states at superficial sites.

Detection of ligand exchange phenomena

The temperature dependence of GCDA and GCA resonances was investigated in the range 280–305 K on samples with a protein/ligand ratio of 1 : 3 (Fig. 3A,B). In both cases, a slow exchange regime on the NMR chemical shift time scale was observed for site 2, which exhibited, upon temperature increase, decreased linewidths, reflecting the shorter protein correlation time at higher temperatures. In contrast, site 1 and the unbound resonances exhibited line broadening upon temperature increase, further confirming the involvement of ligand bound to site 1 in exchange phenomena with the free ligand. Interestingly, at all the investigated temperatures, the resonance of the unbound GCA showed a similar linewidth but a higher intensity with respect to GCDA, reflecting a minor overall affinity of GCA for T91C L-BABP.

One alternative way of detecting exchange phenomena between the different species in solution is through

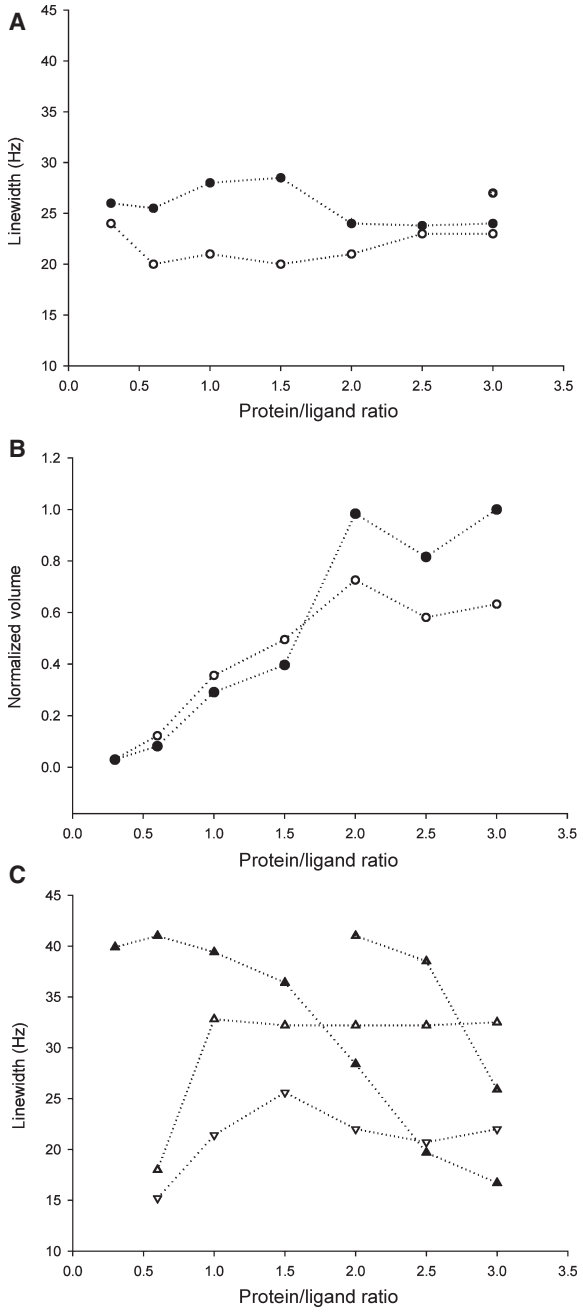


Fig. 2. Analysis of GCA and GCDA resonances at different protein/ligand ratios. Plots of linewidths (A) and volume (B) of amide proton resonances of GCDA as a function of protein/ligand ratio: site 1 (empty circle); site 2 (filled black circle); unbound ligand (filled grey symbols). Plots of linewidths (C) of amide proton resonances of GCA as a function of protein/ligand ratio: site 1 (empty triangle up); site 1′ (empty triangle down); site 2 (filled black triangle); unbound ligand (filled grey triangle).

the measurement of the self-diffusion coefficient (D) [18]. It is expected that a ligand molecule that is in exchange with the free form will show a D-value that is a linear combination of those of the free ligand and the protein. Diffusion experiments were performed on wild-type and T91C L-BABP complexed with GCDA and GCA at protein/ligand ratios of 1 : 3, and the diffusion coefficients, calculated from the analysis of signal decay as a function of the applied gradient, are reported in Table 1. From comparison of these values with those previously obtained for the free ligand (3.97×10^{-6} cm^2·s^{-1}) and the protein (1.04×10^{-6} cm^2·s^{-1}) [14], it is possible to conclude that exchange processes between bound and free forms are relevant for site 1 and negligible for site 2 for both wild-type and T91C L-BABP. However, in the presence of the Cys80–Cys91 disulfide bridge, the diffusion values of ligand bound to site 1 were lower than those of the wild-type protein, suggesting a higher affinity for both ligands at site 1 of T91C L-BABP.

Site selectivity

Previous observations indicated that wild-type L-BABP did not show any site selectivity for GCDA and GCA, and revealed a higher affinity for GCDA at both sites. T91C L-BABP site selectivity for the two bile salts was investigated in competition experiments, in which unlabelled GCDA was added to a solution containing a T91C L-BABP/[^{15}N]GCA molar ratio of 1 : 2. One-dimensional first increments of the 2D ^{1}H/^{15}N correlation spectra for the sample containing an equimolar mixture of [^{15}N]GCA and unlabelled GCDA (Fig. 4A) showed that the peak corresponding to site 2 was sharpened but its intensity was marginally affected by GCDA addition. In contrast, [^{15}N]GCA bound to site 1 was completely displaced by the unlabelled GCDA as its resonance disappeared. This behaviour clearly indicates that the presence of a disulfide bridge had introduced site selectivity. Such an effect was confirmed by the complementary competition experiment, in which the unlabelled GCA was added to a solution containing a T91C L-BABP/[^{15}N]GCDA molar ratio of 1 : 2 (Fig. 4B). In agreement with the selectivity of GCA for site 2, complete disappearance of the resonance of [^{15}N]GCDA at site 2 was expected. However, only a 60% reduction of this resonance intensity was observed, which can be explained by a general overall higher affinity of T91C L-BABP for GCDA than for GCA. Interestingly, the presence of GCA at site 2 favoured one secondary form at a superficial site, characterized by chemical shifts close to the site 1′ resonance, previously observed in ^{1}H/^{15}N-HSQC spectra of the T91C L-BABP–GCDA complex (Fig. 1). The change of the population at site 1 in

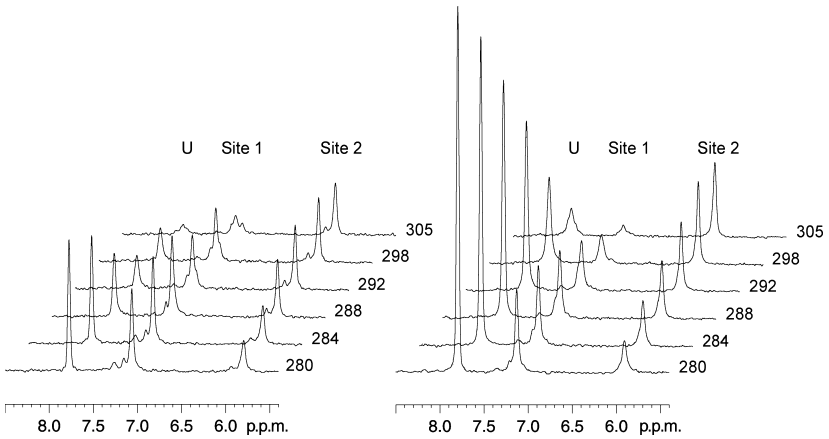


Fig. 3. Stacked plot showing the temperature dependence of the BA amide ^{1}H resonances in the temperature range 280–305 K. One-dimensional first increment of the 2D $^{1}H/^{15}N$-HSQC spectra collected on T91C L-BABP–[^{15}N]GCDA (left) and T91C L-BABP–[^{15}N]GCA (right) complexes, using a protein/ligand molar ratio of 1 : 3.

Table 1. Diffusion coefficients of bile salt species. *D*-values measured for free CDA and holo L-BABP are 3.97×10^{-6} $cm^2 \cdot s^{-1}$ and 1.04×10^{-6} $cm^2 \cdot s^{-1}$, respectively [14]. Errors in *D*-values were estimated to be of the order of 10^{-8} $cm^2 \cdot s^{-1}$ from the fitting procedure.

	Site 1 ($\times 10^{-6}$ $cm^2 \cdot s^{-1}$)	Site 2 ($\times 10^{-6}$ $cm^2 \cdot s^{-1}$)
Wild type–GCDA	2.28	1.57
Wild type–GCA	2.43	1.28
T91C–GCDA	2.02	1.45
T91C–GCA	2.20	1.48

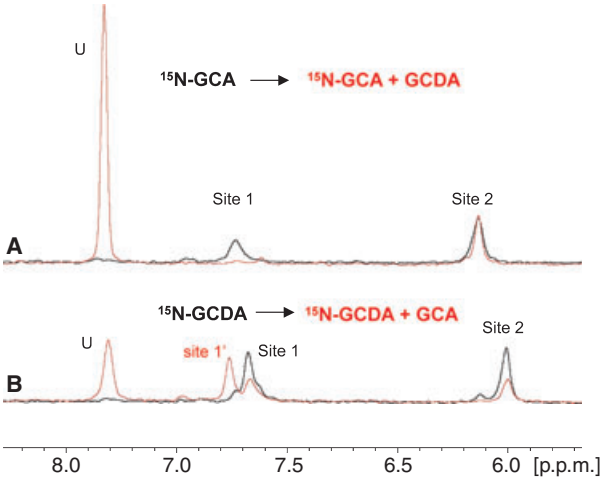


Fig. 4. Bile salt site selectivity experiments. One-dimensional first increment of the 2D $^{1}H/^{15}N$-HSQC spectra collected on: [^{15}N]GCA in a 1 : 2 T91C L-BABP/GCA molar ratio [(A), black line]; [^{15}N]GCA in the presence of equimolar amounts of unlabelled GCDA (T91C L-BABP/GCA/GCDA molar ratio of 1 : 2 : 2) [(A), red line]; [^{15}N]GCDA in a 1 : 2 T91C L-BABP/GCDA molar ratio [(B), black line]; [^{15}N]GCDA in the presence of equimolar amounts of unlabelled GCA (T91C L-BABP/GCA/GCDA molar ratio of 1 : 2 : 2) [(B), red line]. The resonances corresponding to the unbound ligand are indicated as U.

favour of a new site 1′ suggests that site 1′ is the preferred orientation of GCDA at the superficial site when GCA is bound to site 2.

The different intensities exhibited by the unbound GCA and GCDA reflect, as previously observed for wild-type protein, both the lower affinity of the protein for GCA and the onset of different equilibria between monomeric and micellar bile salts. Indeed, the critical micellar concentration of GCDA (2.4 mM) is significantly lower than that of GCA (10 mM) [19], and the broader linewidth of the resonance of unbound GCDA (22 Hz) with respect to that of unbound GCA (15 Hz) can be explained by the equilibrium between free monomeric and micellar GCDA. The comparison of ^{1}H-spectra of the two protein samples at a protein/GCDA/GCA ratio of 1 : 2 : 2 indicated that the final holo state is independent of the order of addition of the bile salts and supports the results of competition data.

In summary, competition experiments pointed to a site preference of GCDA for site 1 and of GCA for site 2 in T91C L-BABP, together with a higher affinity of the protein for GCDA.

Conformational changes induced by disulfide bridge in the apo and holo forms of T91C L-BABP

The effect of the disulfide bond introduction on the structure of the apo protein was investigated by monitoring the $^{1}H/^{15}N$ chemical shifts changes observed in T91C L-BABP with respect to the wild type. The resonance assignment of signals from backbone and side chains atoms of the apo form of T91C L-BABP was performed using standard 3D heteronuclear triple resonance NMR experiments, as described in Experimental

procedures, together with a combination of 2D and 3D TOCSY and NOESY HSQC spectra recorded at pH 5.6 and pH 7.2. The observed shift of some cross-peaks, induced by acidic pH, allowed the assignment of resonances that substantially overlapped at neutral pH. Backbone amide resonance assignment was complete at 93%, and resonances of residues Thr72, Met73, Lys77, Leu78, Asn86, Leu89, Lys95 and Phe96 could not be unequivocally assigned, owing to signal overlap and/or broadening.

The secondary structure of T91C L-BABP is substantially unchanged with respect to the wild-type protein. In particular, the secondary structural elements, as derived with TALOS [20], include 10 antiparallel β-strands and two α-helices in the following regions: 5–8 (strand A), 14–18 (helix 1), 25–29 (helix 2), 34–43 (strand B), 46–53 (strand C), 56–60 (strand D), 66–71 (strand E), 76–85 (strand F), 88–92 (strand G), 96–103 (strand H), 105–113 (strand I), and 116–124 (strand J).

The analysis of chemical shift perturbation (CSP) induced by the introduction of a disulfide bridge showed that the most significant changes occurred at the level of strand E (Ala68, Asp69, and Ile71), strand F (Lys79, Cys80, Thr81, and Leu84), strand G (Ser93), and strand H (His98) (Fig. 5). All of the mentioned residues are in close proximity to the disulfide bridge connecting strand F and strand G, except for Ile71, which is, however, contiguous with the 68–69 region affected by the mutation.

The T91C L-BABP–chenodeoxycholate (CDA) complex was characterized by NMR, and the assignment of backbone amide resonances, performed on a protein/ligand sample of molar ratio 1 : 3, was complete at 95% (missing assignments for Ala1, Gln7, Ile37, Asn86, Gln100, and Asn105). Resonance assignments of apo and holo forms of the protein have been reported in BiomagResBank (accession numbers 16310 and 16309 for the apo and holo proteins, respectively).

Comparison of the chemical shifts of the apo and holo forms of T91C L-BABP indicated that the regions mostly affected by binding are mainly located at the C-terminal part of the protein, at the level of Lys76, Thr81, and Val82 (strand F), Val90, Lys92, and Ser93 (strand G), Glu94 (loop GH), Phe96, Ser97, and His98 (strand H), together with a few residues in the N-terminal region, namely Arg32 (helix II), Thr57 (strand D), and Glu67 (strand E) (Fig. 6). Comparison of CSP induced by complex formation in wild-type and T91C L-BABP (Fig. 6A) indicated that the same protein regions are affected by ligand binding, confirming a conserved binding mode. A few differences were, however, observed for some residues gathered around the ligand bound at site 2 (Fig. 6B), closer to the disulfide bridge.

Backbone dynamics of apo and holo forms of T91C L-BABP

Backbone dynamics were investigated for the apo and holo forms of T91C L-BABP to assess the relevance of backbone motions to ligand-binding properties.

^{15}N T_1 and T_2 relaxation values were calculated for the apo form of T91C L-BABP, and several residues, namely Arg32, Lys52, Phe62, Thr71, Asp74, Cys91, Lys92, Glu94, Ser97, His98, Gln100, Gly104, Glu109, Ile111 and Gly115, showed high T_1/T_2 ratios, indicative of conformational exchange processes on the microsecond and millisecond time scales (Fig. 8). Interestingly, the introduction of the new disulfide bond, connecting strand F and strand G, did not reduce conformational motions, which, on the contrary, were extended to the N-terminal regions of the protein, as a result of changes in motion propagation, within the β-barrel (Fig. S2).

The relaxation experiments were also performed on a holo T91C L-BABP–CDA sample at a protein/ligand ratio of 1 : 3. In these conditions, the protein is substantially saturated and a negligible population of the free protein is present, as derived from the analysis of titration experiments performed on the ^{15}N-labelled protein (data not shown). As a consequence, the detected exchange contribution can be related to protein conformational motions rather than to free-bound exchange. Analysis of T_1/T_2 ratios

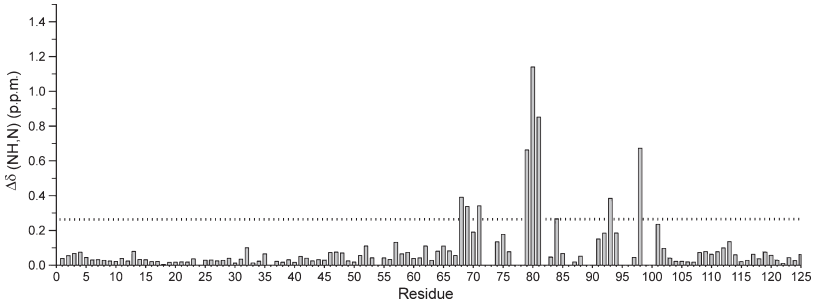
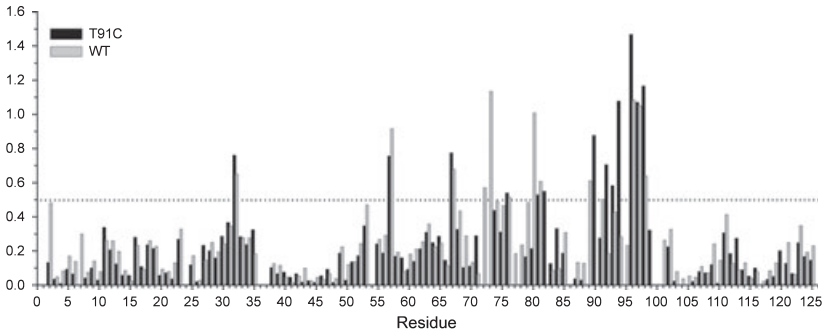


Fig. 5. CSP upon disulfide bridge introduction. Chemical shift differences between apo T91C L-BABP and wild-type (WT) L-BABP, at pH 7 and 298 K, calculated as $\Delta\delta(HN,N) = [(\Delta\delta HN(T91C - WT)^2 + \Delta\delta N(T91C - WT)^2/25)/2]^{1/2}$) are plotted versus residue number. The dotted line corresponds to the mean value plus one standard deviation.



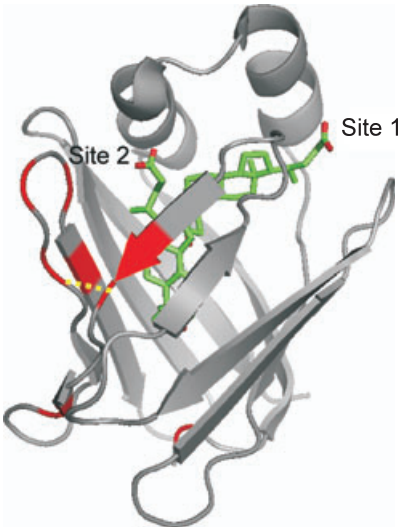


Fig. 6. Chemical shift changes upon CDA binding at pH 7 and 298 K. (A) Chemical shift differences between apo and holo resonances for T91C (black) and wild-type (WT) (grey) L-BABP, calculated as $\Delta\delta(HN,N) = [(\Delta\delta HN(T91C - WT)^2 + \Delta\delta N(-T91C - WT)^2/25)/2]^{1/2}$), are plotted versus residue number. The dotted line corresponds to the mean value plus one standard deviation of T91C L-BABP CSP. (B) Residues showing the major differences upon introduction of a disulfide bridge (Phe2, Lys79, Cys80, Leu84, Lys92, Glu94, Phe96, and His98) are coloured in red on the ribbon representation of L-BABP. The two ligands are coloured in green, and the position of the disulfide bridge is in yellow.

showed that slow motions were not quenched upon ligand binding. Indeed, high T_1/T_2 ratios were observed for Tyr9 and Gln11 (strand A), Arg32 (helix II), Val90, Lys92, and Glu94 (strand G), Phe96 and Ser97 (strand H), Phe113 (strand I), and Arg120 and Val125 (strand J) (Fig. 7). We can conclude that, at variance with what was observed for wild-type protein (Fig. S2), T91C L-BABP complexation with CDA enhanced backbone motions that were already present in the apo protein, except for residues belonging to strand C and strand D and to loop EF and loop IJ.

In view of the physiological relevance of bile salt conjugation, which prevents passive diffusion of bile salts across cell membranes, the NMR analysis was extended to glycoconjugates, namely GCDA and GCA. Both homotypic complexes (T91C

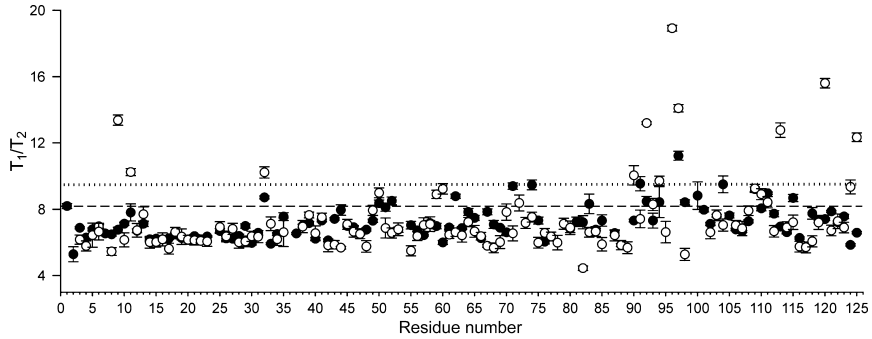


Fig. 7. Comparison of T_1/T_2 ratios for apo and holo T91C L-BABP. [^{15}N]amide T_1/T_2 values as a function of residue number measured at 298 K. Filled black circles: apo T91C L-BABP. Empty circles: T91C L-BABP/CDA at a molar ratio of 1 : 3. Dashed and dotted lines correspond to the mean value plus one standard deviation of apo and holo T91C L-BABP, respectively. Error bars are shown.

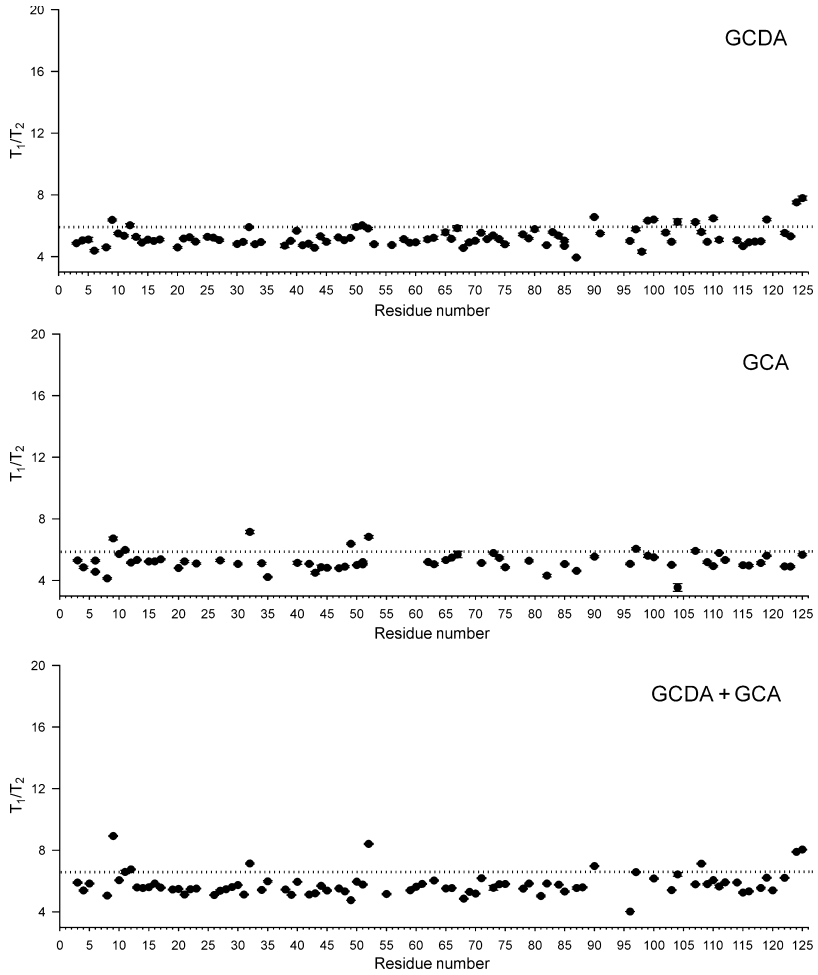


Fig. 8. T_1/T_2 ratios for T91C L-BABP complexed with the different glycoderivatives. [^{15}N]amide T_1/T_2 values as a function of residue number measured at 298 K. Upper panel: T91C L-BABP/GCDA at a molar ratio of 1 : 3. Middle panel: T91C L-BABP/GCA at a molar ratio of 1 : 3. Lower panel: T91C L-BABP/GCDA/GCA at a molar ratio of 1 : 1.5 : 1.5. Dotted lines correspond to the mean value plus one standard deviation of the data. Error bars are shown.

L-BABP/GCDA molar ratio of 1 : 3 and T91C L-BABP/GCA molar ratio of 1 : 3) and the heterotypic complex (T91C L-BABP/GCDA/GCA molar ratio of 1 : 1.5 : 1.5) were characterized according to their relaxation properties. Interestingly, substantial quenching of the motions was observed in the presence of all the glycine derivatives, independent of the hydroxylation pattern (Fig. 8). A few residues at the C-terminal end showed T_1/T_2 ratios higher than one standard deviation for the T91C L-BABP–GCDA complex, whereas the same behavior was observed for residues at the N-terminal end for the T91C L-BABP–GCA complex.

Discussion

Several examples have been reported in the literature, for members of the lipocalin family, where the introduction/removal of a disulfide bridge was responsible for changes in ligand-binding stoichiometry and affinities [11,12,21]. In intracellular proteins, disulfide bonds are generally transiently formed, owing to the reductive nature of the cellular environment. It has been shown that transient disulfide bonds are generally not essential for structural integrity, but can contribute to protein function. Reversible disulfide bridge formation within intracellular proteins can give rise to local and/or global conformational changes that may lead to distinct binding and functional properties [22,23]. In line with this, we have shown here that the presence of a disulfide bridge, while maintaining the same binding stoichiometry, induces changes in binding ability, site selectivity and dynamic properties of L-BABP. Thus, the study of a recombinant protein with a stable disulfide bridge helps in clarifying the role of transient intracellular disulfide bonds.

Both NMR analysis and MS data confirmed the ability of T91C L-BABP to bind two GCDA or GCA molecules, indicating that both protein forms are competent for efficient BA binding and transport within

the cell. These results differ from the recently reported data for the homologous liver zebrafish protein, where the introduction of a disulfide bridge resulted, intriguingly, in a singly ligated protein, with the cholate occupying the more superficial binding site [16].

Exchange peaks observed in $^{1}H/^{15}N$-HSQC spectra of holo proteins, together with diffusion experiments, showed that exchange processes between bound and free forms are relevant for site 1 and negligible for site 2, independently of the presence of a disulfide bridge (Table 1). The introduction of a disulfide bridge induced significant changes in the GCA exchange regime for ligand bound to site 1, whose resonance was observable at all the investigated protein/ligand ratios, at variance with the wild-type protein [5]. In line with this observation is the trend of diffusion coefficients measured for GCA bound to T91C and wild-type L-BABP, pointing to a higher affinity of this ligand for T91C L-BABP site 1 (Table 1).

The most relevant feature emerging from the analysis presented here is the ability of the disulfide bridge to modulate recognition at both sites. Indeed, no site selectivity was previously observed for wild-type L-BABP [5], whereas it is now clear that when T91C L-BABP is incubated with only GCDA or GCA, both binding sites are occupied, but when the two bile salts, differing only in hydroxylation at position 12, are present, GCDA preferentially binds to site 1 and GCA to site 2. Site selectivity is, however, observed only when both GCDA and GCA are present, suggesting that it does not derive from steric exclusion of one bile salt from a specific site.

Protein observation was required in order to investigate the structural basis of these varied ligand-binding properties. Both CSP (Fig. 5) and TALOS analysis on the apo protein indicated that no significant change in 3D structure occurred. The comparison of CSP for the holo forms of T91C L-BABP and the wild type (Fig. 6A) indicated that the same protein regions are generally involved in ligand binding, even if all the residues showing significantly different CSP values in the two proteins were gathered around the ligand bound at site 2 (Fig. 6B). This result is in perfect agreement with data derived from ligand observation (Fig. 9), revealing significant changes in the chemical shifts of site 2 resonance for both bile salts. This behaviour is ascribed to local changes in the chemical environment due to the introduction of a disulfide bridge, which involves two residues that are in contact with the ligand bound to the 'internal' binding site in the holo wild-type structure (Protein Data Bank ID: 2JN3 [14]).

Protein dynamics is largely influenced by the presence/absence of the disulfide bridge. Indeed, the presence of the disulfide bridge favoured the propagation of slow motions from the C-terminal region of the molecule to the N-terminal β-sheet in the apo protein, and enhanced backbone motions in the T91C L-BABP–CDA complex, at variance with the behaviour of the wild-type protein, where the binding of this ligand was accompanied by substantial quenching of motions (Fig. S2).

Molecular dynamics simulation studies revealed differently coupled correlated motions for some iLBPs, depending on the presence and the type of ligand [10,24]. These data prompted us to evaluate the effect of BA glycosylation and hydroxylation pattern on protein conformational motions. Interestingly, all glycoderivative mixtures were efficient in reducing backbone dynamics (Fig. 8), possibly as a consequence of the onset of more favourable interactions between the glycine moiety and the protein portal region. Indeed, comparison of the CSP in the presence of CDA or GCDA (Fig. 10) suggested that the most affected

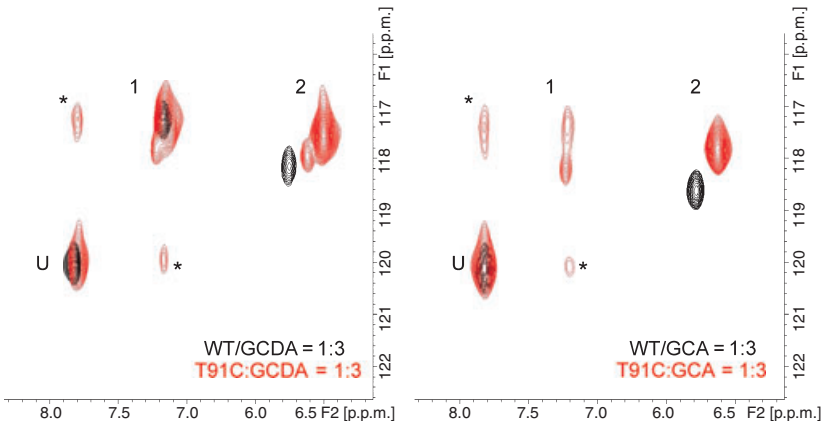
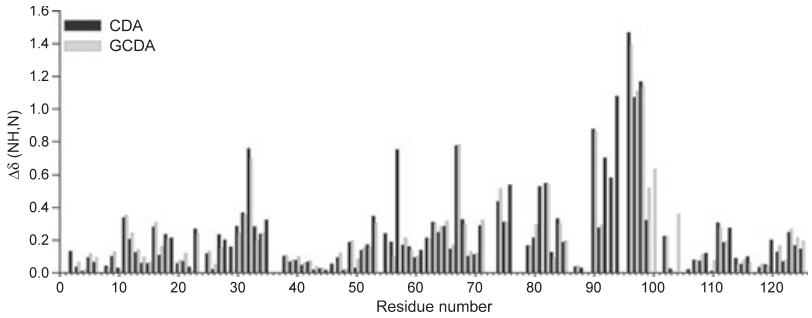


Fig. 9. Comparison of [^{15}N]GCDA and [^{15}N]GCA in complex with wiild-type (WT) L-BABP (black) and T91C L-BABP (red). Superposition of 2D $^{1}H/^{15}N$-HSQC spectra of [^{15}N]GCDA (left panel) and [^{15}N]GCA (right panel) at a 1 : 3 protein/ligand molar ratio. The resonances corresponding to the unbound ligand and to binding sites 1 and 2 are indicated as U, 1, and 2, respectively. Exchange peaks between site 1 and unbound resonance are labelled with asterisks.



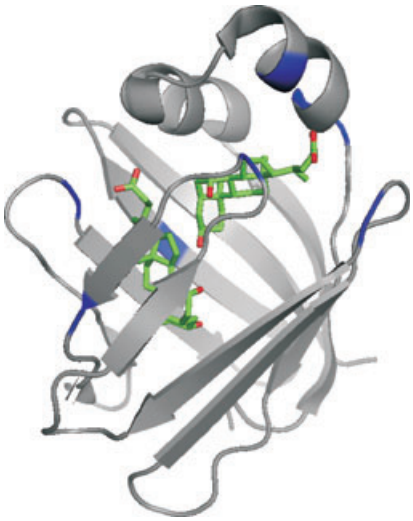


Fig. 10. Chemical shift changes upon CDA or GCDA binding to T91C L-BABP at pH 7 and 298 K. (A) Chemical shift differences between apo and holo resonances for CDA (black) and GCDA (grey) complexes, calculated as $\Delta\delta(HN,N) = [(\Delta\delta HN(T91C - WT)^2 + \Delta\delta N(T91C - WT)^2/25)/2]^{1/2}$), are plotted versus residue number. The dotted line corresponds to the mean value plus one standard deviation of T91C L-BABP CSP. (B) Residues showing the major differences in the two complexes (Tyr9, Leu27, Gln42, Val49, Thr50, Thr59, Asp74, Cys80, Lys86, and Arg124) are coloured in blue on a ribbon representation of L-BABP.

residues are located at the level of the portal area, as expected in response to the protrusion of the glycine moieties, and at the level of strand F, strand H and strand I, in close contact with the ligand bound at site 2. Specifically, the chemical shift variation observed at the portal area for Arg32 and Asp33 suggests a different positioning of helix II in the two complexes. Arg32, characterized by high T_1/T_2 values in the apo protein and in all of the investigated holo proteins, thus plays a key role in regulating the positioning of the helix–loop–helix motif with respect to the β-barrel in order to accommodate the different BAs.

Analysis of relaxation data obtained for the glycoderivatives showed that GCDA was able to quench motions affecting the protein open end (helical and loop EF regions), whereas the bound GCA mostly influenced the C-terminal region of the protein, in agreement with the site selectivity observed for the two ligands. Interestingly, the heterotypic complex, in which the proper ligand is expected to be located at the corresponding binding site, still presented a few residues with high T_1/T_2 values, especially at the N-terminal end. The competition data (Fig. 4) indicated that GCDA preferentially populates site 1′ when GCA is bound to site 2, and this different orientation at the superficial site may induce different motional properties at the N-terminal end. The detected site preferences and changes in chemical shifts in heterotypic complexes further establish a parallelism with the behaviour observed for I-BABP and its mutants [8], thus substantiating the previous proposal that BABPs exert a parallel function in hepatocytes and enterocytes [4,25].

In conclusion, it is shown here that the introduction of a disulfide bond makes the protein competent for site selectivity. NMR data indicated that protein conformational changes induced by the disulfide bond are small and gathered around the inner binding site, whereas significant changes in backbone motions are observed for several residues distributed over the entire protein. Site selectivity appears, therefore, to be governed by protein mobility, rather than by steric factors related to the hydroxylation pattern of the ligand, in agreement with what has been observed for other BABPs [4,25]. These results once more underline the tight connection between ligand-binding phenomena

and protein mobility in this protein family, and set the basis for further NMR kinetic studies based on lineshape analysis and relaxation dispersion measurements.

Experimental procedures

Protein expression and purification

The expression plasmid for T91C L-BABP was obtained from that of wild-type L-BABP using the Quickchange (Stratagene, La Jolla, CA, USA) mutagenesis kit. The presence of the desired mutation was confirmed by plasmid sequencing. Recombinant T91C L-BABP was expressed from *Escherichia coli* and purified to homogeneity as previously described [9]. Delipidated T91C L-BABP was obtained in a yield of 95 mg·L^{-1} of rich medium. ^{13}C/^{15}N labelling was achieved using M9 minimal media containing 1 g·L^{-1} $^{15}NH_4Cl$ and 4 g·L^{-1} ^{13}C-enriched glucose, following protocols reported in the literature [26]. ^{15}N-labelled and ^{13}C/^{15}N-labelled T91C L-BABP were both obtained with a 75 mg·L^{-1} yield from minimal media. Protein concentrations for sample preparation were determined spectrophotometrically. The presence of the disulfide bridge was confirmed by MS.

NMR sample preparation

NMR studies on the apo protein were performed on 0.5 mM ^{15}N/^{13}C-labelled samples of T91C L-BABP dissolved in 30 mM potassium phosphate buffer in 95% H_2O/5% D_2O. The pH of the solutions was 5.6 or 7.2. Unenriched BAs and [24-^{13}C]glycocholate were purchased from Sigma (St Louis, MO, USA). [^{15}N]Glycine conjugates of CDA and CA were prepared as previously reported [5]. The titration of the unlabelled T91C L-BABP with increasing amounts of [^{15}N]GCDA or [^{15}N]GCA was performed at seven protein/ligand ratios (1 : 0.3, 1 : 0.6, 1 : 1, 1 : 1.5, 1 : 2, 1 : 2.5, and 1 : 3), and the preparation of the holo protein samples was performed following a procedure previously described [5]. Protein–ligand complexes were analysed at pH 7.2 on 0.5 mM T91C L-BABP samples; each protein/ligand molar ratio sample was prepared and analysed twice, in order to minimize errors.

NMR data collection and analysis

NMR spectra were acquired at 298 K on Bruker DMX 500 and Avance III 600 spectrometers equipped with a 5 mm TCI cryoprobe and a *Z*-field gradient. Data were processed with NMRPIPE [27] and visualized with NMRVIEW [28]. For the assignment of apo and holo protein resonances, the following experiments, ^{1}H/^{15}N-TOCSY HSQC and NOESY HSQC, together with HNCACB, CBCA(CO)NH and HN(CA)CO were performed. The secondary structure elements were derived with the software TALOS [20] from chemical shift data of HN, N, HA, CA and CB nuclei.

^{15}N-relaxation experiments

^{15}N-relaxation experiments for apo and holo (T91C L-BABP/CDA molar ratio of 1 : 3) samples were acquired at 600 MHz at pH 7.2. A dataset of 14 variable delays (2.5, 20, 60, 100, 150, 200, 300, 400, 600, 800, 1000, 1500, 1700 and 2500 ms) was used for T_1 measurements, and a dataset of nine variable delays (16.96, 33.92, 50.88, 67.84, 101.76, 135.68, 169.6, 220.48 and 237.44 ms) was used for T_2 measurements. For T91C L-BABP in complex with CDA, a dataset of nine variable delays (0.01, 180, 360, 540, 720, 900, 300, 1080, 1260 and 1440 ms) was used for T_1 measurements, and a dataset of seven variable delays (16.96, 33.92, 50.88, 67.84, 101.76, 220.48 and 237.44 ms) was used for T_2 measurements. T_1 and T_2 values were determined for 112 nonoverlapping cross-peaks. For the holo T91C L-BABP in complex with GCDA, GCA, or both (T91C L-BABP/GCDA/GCA molar ratio of 1 : 1.5 : 1.5), 10 variable delays (10, 60, 180, 300, 450, 600, 740, 900, 1100, 1200 and 1400 ms) were used for T_1 measurements, and 10 variable delays (16.98, 33.16, 49.74, 66.32, 82.9, 99.48, 132.64, 149.22, 198.96 and 232.12 ms) were used for T_2 measurements, recorded at 500 MHz. T_1 and T_2 relaxation values were estimated for 92, 65 and 86 residues for the complex with GCDA, the complex with GCA, and the heterotypic complex, respectively.

Titration experiments

^{1}H/^{15}N-HSQC spectra of unlabelled protein complexed with labelled ligands were acquired with a ^{1}H spectral width of 6510 Hz and 1024 points, zero-filled to a total of 2048 points. Relaxation delays of 1.7 s were employed. In the ^{15}N dimension, 256 increments were collected, with a sweep width of 2032 Hz, zero-filled to a total of 1024 points.

The linewidth dependence of ligand ^{1}H resonances as a function of protein/ligand ratio was followed through the first increment of 2D ^{1}H/^{15}N-HSQC spectra recorded under identical conditions (8000 points on a sweep width of 6510 Hz).

The temperature dependence of the BA amide ^{1}H resonances for a protein/ligand molar ratio of 1 : 3 was followed through the first increment of a 2D ^{1}H/^{15}N-HSQC spectrum collected with 8000 points on a sweep width of 6510 Hz in the temperature range 280–305 K.

Diffusion experiments

^{15}N-edited diffusion experiments were performed on samples of wild-type and T91C L-BABP in complex with

[^{15}N]GCDA and [^{15}N]GCA at protein/ligand ratios of 1 : 3, in order to determine the diffusion coefficients of protein-bound ligands as compared with those of the free molecules. The pulse program was obtained by combining the standard HSQC pulse scheme with a pulsed-field gradient stimulated echo module employing bipolar gradients under the same conditions previously reported [6]. The measured signal volumes as a function of the applied gradient were fitted to the following equation, using a nonlinear least squares minimization:

$$I = I(0)\exp[-D\gamma^2G^2\delta^2(\Delta - \delta/3 - \tau/2) \quad (1)$$

where D is the translational diffusion coefficient, γ is the ^{1}H gyromagnetic ratio, G is the gradient strength, Δ and δ are as defined above, and τ is the gradient pulse separation.

Acknowledgements

S. Zanzoni and M. Guariento are gratefully acknowledged for help in protein expression and purification. We are grateful to R. Longhi for providing the BA [^{15}N]glycine conjugates. This research was supported by FIRB 2003 (Project No. RBNE03PX83), Cariverona Foundation. The University of Verona is acknowledged for financial support in the acquisition of the NMR Bruker Avance 600 MHz spectrometer equipped with a cryoprobe. L. Ragona thanks CNR-RSTL 2007 (Code No. 779) for financial support. C. Cogliati was supported by a grant 'Sovvenzione Globale INGENIO' from 'Fondo Sociale Europeo, Ministero del Lavoro e della Previdenza Sociale and Regione Lombardia'. CIRMMP (Consorzio Interuniversitario di Risonanze Magnetiche di Metalloproteine Paramagnetiche) is gratefully acknowledged.

References

1 Houten SM, Watanabe M & Auwerx J (2006) Endocrine functions of bile acids. *EMBO J* **25**, 1419–1425.
2 Chawla A, Saez E & Evans RM (2000) Don't know much bile-ology. *Cell* **103**, 1–4.
3 Thomas C, Pellicciari R, Pruzanski M, Auwerx J & Schoonjans K (2008) Targeting bile-acid signalling for metabolic diseases. *Nat Rev Drug Discov* **7**, 678–693.
4 Tochtrop GP, DeKoster GT, Covey DF & Cistola DP (2004) A single hydroxyl group governs ligand site selectivity in human ileal bile acid binding protein. *J Am Chem Soc* **126**, 11024–11029.
5 Tomaselli S, Ragona L, Zetta L, Assfalg M, Ferranti P, Longhi R, Bonvin AM & Molinari H (2007) NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids. *Proteins* **69**, 177–191.
6 Pedò M, D'Onofrio M, Ferranti P, Molinari H & Assfalg M (2009) Towards the elucidation of molecular determinants of cooperativity in the liver bile acid binding protein. *Proteins Struct Funct Bioinformatics* doi:10.1002/prot.22496.
7 Tochtrop GP, Bruns JL, Tang C, Covey DF & Cistola DP (2003) Steroid ring hydroxylation patterns govern cooperativity in human bile acid binding protein. *Biochemistry* **42**, 11561–11567.
8 Toke O, Monsey JD, DeKoster GT, Tochtrop GP, Tang C & Cistola DP (2006) Determinants of cooperativity and site selectivity in human ileal bile acid binding protein. *Biochemistry* **45**, 727–737.
9 Ragona L, Catalano M, Luppi M, Cicero D, Eliseo T, Foote J, Fogolari F, Zetta L & Molinari H (2006) NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J Biol Chem* **281**, 9697–9709.
10 Eberini I, Guerini Rocco A, Ientile AR, Baptista AM, Gianazza E, Tomaselli S, Molinari H & Ragona L (2008) Conformational and dynamics changes induced by bile acids binding to chicken liver bile acid binding protein. *Proteins* **71**, 1889–1898.
11 Liu J, Chenyun G, Yihe Y & Donghai L (2008) Effects of removing a conserved disulfide bond on the biological characteristics of rat lipocalin-type prostaglandin D synthase. *Biochimie* **90**, 1637–1646.
12 Sjoelund V & Kaltashov IA (2007) Transporter-to-trap conversion: a disulfide bond formation in cellular retinoic acid binding protein I mutant triggered by retinoic acid binding irreversibly locks the ligand inside the protein. *Biochemistry* **46**, 13382–13390.
13 Vasile F, Ragona L, Catalano M, Zetta L, Perduca M, Monaco H & Molinari H (2003) Solution structure of chicken liver basic fatty acid binding protein. *J Biomol NMR* **25**, 157–160.
14 Eliseo T, Ragona L, Catalano M, Assfalg M, Paci M, Zetta L, Molinari H & Cicero DO (2007) Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two bile salts revealed by NMR. *Biochemistry* **46**, 12557–12567.
15 Nichesola D, Perduca M, Capaldi S, Carrizo ME, Righetti PG & Monaco HL (2004) Crystal structure of chicken liver basic fatty acid-binding protein complexed with cholic acid. *Biochemistry* **43**, 14072–14079.
16 Capaldi S, Guariento M, Saccomani G, Fessas D, Perduca M & Monaco HL (2007) A single amino acid mutation in zebrafish (Danio rerio) liver bile acid-binding protein can change the stoichiometry of ligand binding. *J Biol Chem* **282**, 31008–31018.
17 Reibarkh M, Malia TJ & Wagner G (2006) NMR distinction of single- and multiple-mode binding of small-molecule protein ligands. *J Am Chem Soc* **128**, 2160–2161.

18 Hsu EW & Mori S (1995) Analytical expressions for the NMR apparent diffusion coefficients in an anisotropic system and a simplified method for determining fiber orientation. *Magn Reson Med* **34**, 194–200.
19 Nakashima T (2002) Potentiometric study on critical micellization concentrations (CMC) of sodium salts of bile acids and their amino acid derivatives. *Colloids Surf B Biointerfaces* **24**, 103–110.
20 Cornilescu G, Delaglio F & Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**, 289–302.
21 Capaldi S, Perduca M, Faggion B, Carrizo ME, Tava A, Ragona L & Monaco HL (2007) Crystal structure of the anticarcinogenic Bowman–Birk inhibitor from snail medic (Medicago scutellata) seeds complexed with bovine trypsin. *J Struct Biol* **158**, 71–79.
22 Piotukh K, Kosslick D, Zimmermann J, Krause E & Freund C (2007) Reversible disulfide bond formation of intracellular proteins probed by NMR spectroscopy. *Free Radic Biol Med* **43**, 1263–1270.
23 Thangudu RR, Manoharan M, Srinivasan N, Cadet F, Sowdhamini R & Offmann B (2008) Analysis on conservation of disulphide bonds and their structural features in homologous protein domain families. *BMC Struct Biol* **8**, 55.
24 Woolf TB, Grossfield A & Tychko M (2000) Differences between apo and three holo forms of the intestinal fatty acid binding protein seen by molecular dynamics computer calculations. *Biophys J* **78**, 608–625.
25 Guariento M, Raimondo D, Assfalg M, Zanzoni S, Pesente P, Ragona L, Tramontano A & Molinari H (2008) Identification and functional characterization of the bile acid transport proteins in non-mammalian ileum and mammalian liver. *Proteins* **70**, 462–472.
26 Marley J, Lu M & Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20**, 71–75.
27 Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.
28 Johnson B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313–352.

Supporting information

The following supplementary material is available:
Fig. S1. Chenodeoxycholic acid structure.
Fig. S2. Comparison of T_1/T_2 ratio for apo and holo T91C and wild-type proteins.

This supplementary material can be found in the online version of this article.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

7. NMR lineshape and relaxation analysis of the binding mechanism of cL-BABPs with two bile acids

Lineshape analysis of NMR signals and relaxation dispersion experiments can provide a direct view of the mechanism and the kinetic rates of binding events or exchange processes. In the following paper [*Chemistry, A European Journal*, submitted], lineshape analysis of two-dimensional ^{1}H-^{15}N-HSQC spectra and relaxation dispersion experiments have been used to probe the molecular recognition and the complex uptake mechanism of two bile salt molecules by BABPs.

- cL-BABP/S-S and cL-BABP were titrated with GCDA. For each of the thirteen steps of the titration, in the protein:ligand ratios range 1:0 to 1:3.5, a ^{1}H-^{15}N-HSQC spectrum was recorded.
- NMR lineshape analysis of individual residues provided a direct view of the interaction mechanisms. A "conformational selection" model has been postulated as the initial recognition process, where the dynamics observed in the apo-form is essential for ligand uptake, leading to conformations with improved access to the cavity.
- Relaxation dispersion experiments were further performed to investigate the nature of interconverting conformers in the apo form. The analysis of the obtained data confirmed the existence of a conformational selection model, and the presence of a "holo like" conformation in the absence of the ligand.
- Subsequent multistep events could be modelled, in few instances, with a two step binding mechanism, where intermediate species should be considered even if they were not always easily detectable. Nature, position and lifetime of the intermediates could be influenced by several factors and could correspond to various conformations adopted by the protein in its free state or bound with different amount of GCDA.
- Both cL-BABP/S-S and cL-BABP, in the ligand-bound states, exhibit a slow time-scale flexibility, suggesting a reorganisation of the protein occurring after that saturation is reached.
- A common binding mechanism was proposed for cL-BABP/S-S and cL-BABP on the basis of lineshapes analysis.

- The main differences between the two proteins behaviour are related to their affinity for GCDA, influencing the kinetics of the first exchange process.

My contribution to the following paper, titled "Site-specific investigation of steady-state kinetics and dynamics of the multistep binding of bile acid molecules to a lipid carrier protein", consisted of:

1. A period of work at the NMR laboratory of the University of Birmingham, under the supervision of Prof. U.L. Günther, (i) to acquire the knowledge necessary to deal with lineshape and relaxation dispersion studies, through an appropriate analysis and simulation procedures, using the dedicated NMRKIN software, and (ii) to register NMR titration experiments necessary for the study of the protein-ligand interaction mechanisms;
2. Setting the NMR experiments for cL-BABPs titration, finding conditions and parameters to better understand the interaction mechanisms;
3. Acquisition and analysis of the obtained data;
4. Paper writing in which I prepared the result section and significantly contributed to the development of the discussion.

DOI: 10.1002/cbic.200((will be filled in by the editorial staff))

Kinetic mechanism of bile salt binding to a lipid carrier

Site-specific investigation of steady-state kinetics and dynamics of the multistep binding of bile acid molecules to a lipid carrier protein.

Clelia Cogliati[a,b], Laura Ragona[a], Serena Zanzoni[b], Ulrich Günther[c], Sara Whittaker[c], Christian Ludwig[c], Simona Tomaselli[a], Michael Assfalg[b] and Henriette Molinari[b]**

The investigation of multi-site ligand protein binding and multistep mechanisms is highly demanding and represents a fundamental question at the interface between chemistry and biology. Advanced NMR approaches, such as 2D ^{1}H-^{15}N line shape analysis, allowing a reliable investigation of ligand binding occurring on microsecond to millisecond time scales, have been here extended to model, for the first time, a two-step binding mechanism. The molecular recognition and the complex uptake mechanism of two bile salt molecules by lipid carriers is an interesting example, here discussed, showing that protein dynamics has the potential to modulate the macromolecule-ligand encounter. Kinetic analysis supports a conformational selection model as the initial recognition process, where the dynamics observed in the apo-form is essential for ligand uptake, leading to conformations with improved access to the cavity. Subsequent multistep events could be modelled, for several residues, with a two-step binding mechanism. The protein, in the ligand-bound state, still exhibits a conformational rearrangement occurring on a very slow time-scale, as observed for other proteins of the family. A global mechanism suggesting how bile acids accesses the macromolecular cavity is thus proposed.

Introduction

It is now well recognized that bile acids are important signalling molecules that regulate a network of metabolic pathways including lipid, glucose, drug and energy metabolism[1, 2], although the specific mechanisms that couple intracellular lipid processing to biological targets and signalling pathways are not yet well understood. In recent years, a number of transport proteins expressed in the liver and intestine, belonging to the intracellular lipid binding protein family (iLBP), and specifically involved in bile acid transport, have been recognized to play central roles in driving bile flow, with complex regulation of activity and function in the nucleus, cytoplasm and membrane.

Most members of iLBPs bind a single fatty acid or retinoid molecule with relatively high affinity[3-5]. Instead, proteins belonging to the bile acid binding protein (BABP) subfamily, seem to bind two or more ligands, with differences in ligand selectivity, binding affinity, stoichiometry and binding mechanism [6-8]. This remarkable binding behaviour has implications for the biological function and could be relevant for regulating the passage of bile acids through the cytoplasm.

The investigation of multi-site ligand protein binding and of multi-step mechanisms is highly demanding and a variety of biophysical techniques have been applied to tackle this problem. New experimental approaches, provided by advances in nuclear magnetic resonance (NMR) spectroscopy, have been described which allow a reliable and sensitive investigation of ligand binding events occurring on microsecond to millisecond (µs-ms) time scales using line shape analysis and relaxation dispersion experiments [9]. Particularly, the combination of these methods can be useful in the study of complex mechanisms, allowing the correlation between protein internal dynamics and function[10]. However applications have been restricted, up to now, to binary systems, such as, among others, retinol interaction with cellular retinol binding protein[9] belonging to the here described iLBP family.

The investigation of the binding mechanism leading to the formation of ternary complexes, in systems endowed with different levels of cooperativity, such as those formed by bile acid binding proteins and their lipidic ligands[11]; [12]is a complex task which has been addressed, for example, through stopped-flow fluorescence kinetic measurements in the case of human and rabbit BABPs[13, 14]. Here we report an investigation, employing

[a] *C. Cogliati, L. Ragona, S. Tomaselli*
Laboratorio NMR
ISMAC-CNR
Via Bassini 15, 20133 Milano, Italy
Fax: +390223699620
E-mail: laura.ragona@ismac.cnr.it
[b] *S. Zanzoni, M. Assfalg, H. Molinari*
NMR Laboratory, Biotechnology Department,
University of Verona
Strada Le Grazie 15, 37134 Verona, Italy
Fax: +390458027929
E-mail: henriette.molinari@univr.it
[c] *U. Günther, S. Whittaker, Christian Ludwig*
School of Cancer Sciences
University of Birmingham
Vincent Drive, Birmingham, B152TT, UK
Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

NMR lineshape analysis and relaxation dispersion experiments, on the multistep binding mechanisms at work in the formation of ternary complexes between chicken liver BABPs and two molecules of sodium glycochenodeoxycholate (GCDA), the most abundant salt of the bile acid pool. Chicken liver BABP displays the typical barrel structure of the family, formed by two orthogonal β–sheets, and a helix–loop–helix motif defining, with flexible loops, the so-called protein open end, delimiting the entrance to the cavity hosting lipophilic ligands. The two most extensively characterized BABPs, human ileal BABP (I-BABP)[11, 15] and chicken liver BABP (L-BABP)[15], share the common property of binding two bile salt molecules with weak intrinsic affinities and strong positive cooperativity. A number of factors have been shown to modulate ligand binding, e.g. the chemistry of the ligand, the nature of the protein residues[6, 16] and the presence of disulfide bridges[17]. For chicken L-BABP two forms are known, differing by the presence of a cysteine or a threonine at position 91 of the primary sequence, resulting in the formation of a protein with (L-BABP/S-S) or devoid of an intramolecular disulphide bridge (L-BABP), which was shown to play an important role in regulating site selectivity[17]. Although the ligand binding mechanism for both L-BABP/S-S and L-BABP, appears to be too complex to determine the rates for all the individual steps, the overall mechanism is well defined. The data presented allow the distinction between two different mechanisms in response to bile acid binding, namely induced fit versus conformational selection. This analysis provides strong evidence that molecular recognition is promoted by a conformational selection process, where the protein dynamics improves access to the cavity and is therefore essential for ligand uptake. Subsequent multistep events could be modelled, in few instances, with a two-step binding mechanism, providing information on the involved kinetic parameters. Finally, the protein populates an ensemble of states, even in the ligand-bound state.

Results and Discussion

Line shape analysis of NMR resonances can provide a direct view of the mechanisms and the kinetic rates of binding events or exchange processes[18]. It has been reported[19] that this approach can also be applied to protein signals in 2D NMR spectra of ^{15}N labelled proteins, allowing kinetic information to be obtained for individual residues of the protein. In the case of protein ligand interactions the kinetic mechanism can be studied by recording spectra at different concentrations of the binding partner. Although experimental data often exhibit complex behaviours, it was shown[18] that the analysis of theoretical line shapes derived for different simplified mechanisms allows the identification of the type of kinetic mechanism responsible for a particular behaviour under steady-state conditions.

In this study line shape analysis of NMR signals has been used to elucidate the mechanism of bile salt binding to L-BABPs.

Thirteen ^{1}H-^{15}N HSQC spectra obtained for different protein:ligand (P:L) ratios in the range 1:0 to 1:3.5, were analysed. The highest P:L ratio was chosen to exceed the reported 1:2 stoichiometry, to ensure complete protein saturation. 1D cross-sections, derived from these two-dimensional NMR spectra, at the ^{1}H and ^{15}N resonance frequencies of individual protein residues provide information on the effect of binding on two adjacent nuclei along the entire titration pathway. The two neighbouring nuclei should experience the same kinetic mechanism, although differential effects could arise from a lack of resolution in one dimension. Owing to resonance overlap, the detailed kinetic evaluation could only be performed for sixty-two residues which are evenly distributed along the protein sequence, and are thus representative of all regions of the protein.

In order to correctly monitor the behaviour of individual cross-peak resonances during the titration, backbone resonances of holo L-BABP/S-S complexed with GCDA at a P:L ratio 1:2 were fully assigned, employing standard 3D heteronuclear experiments. The assignment is reported in Table S1 in Supporting information.

Line shapes analysis for steady-state spectra along the titration path of L-BABP/S-S reveals different behaviours, as summarised in Table 1. The first observation derived from the cross-sections analysis is that some residues (Table 1), mostly distributed on the β-barrel, are characterised by a low intensity signal for the apo form, suggesting conformational exchange before ligand addition and/or exchange with solvent (Figure 1a). However, most of these residues are involved in H-bonds, as evidenced both by H/D exchange experiments[20] and by the analysis of the apo X-ray structure (PDB code: 1tvq), suggesting that the observed reduced intensity mainly derives from conformational heterogeneity. Residues showing this type of behaviour form a continuous ridge crossing the protein from the loop between the two helices to the C strand, involving residues in the EF loop, E and D strands (Figure 1b), thus suggesting a possible path for correlated motions. It is therefore plausible that the observed conformational averaging provides a conformational space favouring ligand binding, i.e. an equilibrium P ⇄ P*, where P* represents a conformation competent for ligand uptake into the internal protein cavity.

Relaxation dispersion (RD) experiments were further performed to investigate the nature of the interconverting conformers in the absence of ligand. RD experiments detect dynamics on a similar time scale as line shape analysis and can also provide information about invisible states with populations >0.5% in exchange with a more highly populated ground state[21]. This analysis was performed for 71 residues of apo L-BABP/S-S, free of overlap, with RD spectra recorded at 600 and 900 MHz. 48 residues showed a contribution of R_{ex} to the transverse relaxation rate R_2, and dispersion profiles like those reported in Figure 2a were obtained. The kinetics of the exchange process, the populations of the exchanging states and the absolute values of chemical shift differences between ground and excited states were extracted, assuming a simple two-state exchange model. The chemical-shift differences obtained from relaxation dispersion data ($\Delta\omega$) range between 0.3 and 2.6 ppm. The dynamic regions are evenly distributed in the protein scaffold (Figure 2b) and map those regions that show the largest chemical shift changes upon bile acid binding (Figure 2c), suggesting that the free protein samples a minor conformation similar to the ligand-bound state. The chemical shift changes between major and minor species obtained from relaxation dispersion data ($\Delta\omega$) potentially provide structural information on otherwise inaccessible conformational states. Linear correlations of $\Delta\omega$ *vs.* the chemical shift differences between apo and holo states ($\Delta\delta$) have been interpreted as a small population of a high energy conformation similar to the final holo state arising from the apo protein[22]. Such a behaviour, when observed, is an indication that a "conformational selection" mechanism is driving the ligand recognition process.

For L-BABP/S-S in spite of a generally good qualitative agreement between RD results ($\Delta\omega$) and the chemical shift differences between apo and holo states ($\Delta\delta$) (Figure 2b, c), a limited quantitative agreement was observed. The observed differences were larger than the experimental uncertainties, despite the good fit obtained for the RD spectra. The lack of correlation can be explained in part by the presence of direct interactions between several residues of L-BABP/S-S and the ligand, which are expected to have a significant influence on the chemical shifts of the bound state. Furthermore differences in chemical shift changes determined by RD and HSQC of different states of a protein are often indicative of a more complex binding mechanism involving states with chemical shifts different to those of the apo and holo forms[23]. Thus RD data may not be directly reconciled within the framework of the two-state exchange model employed for the fitting, for example if there is a larger number of conformers in exchange with the apo form. Nevertheless, for a subgroup of residues we observe a good correlation between $\Delta\omega$ and $\Delta\delta$ (Figure 2d). These residues are mainly located in three protein regions, namely the helical and loop regions defining the protein open end, the central beta-sheet and the protein anti-portal region (Figure 2e). A mechanism for ligand uptake, termed the "portal hypothesis" was previously proposed where the

natural ligand enters the protein through a dynamic area, consisting of the helical region and the turns between CD and EF strands, before binding inside the cavity[24]. The behaviour of BABP/S-S "portal residues" is consistent with the hypothesis that a P ⇄ P* equilibrium involves a holo-like conformation, where the portal region visits a ligand-accessible open conformation, thus supporting a "conformational selection" model of binding.

Line shape simulation was then applied, where possible, to analyse ^{1}H and ^{15}N cross-sections at each titration point. A simple one step-model was initially employed but then discarded as simulated lines differed from those observed for most of the analyzed residues. Therefore a kinetic model describing a two-step binding mechanism was implemented in the NMRKIN software (see Experimental session)[10].for the protein in complex with two ligands. The quantitative analysis yielding kinetic parameters and a validation of this kinetic model was feasible for a group of twenty-one residues (see group A, Table 1).The two-step binding mechanism can be described by a general equation of the type:

$$P^*+L \rightleftarrows PL_n+L \rightleftarrows PL_1L_2 \quad [1]$$

where PL_n represents a generic intermediate which could derive either from differences in protein conformations or in the type of bound species. Such an intermediate is generally not directly observed owing to fast exchange with apo or holo protein and/or small populations of intermediate states. In some cases, however, shoulders in HSQC cross-sections clearly indicated the presence of further intermediates. Simulations required that both steps are ligand dependent, i.e. $k_{on}=k_{on}^*[L]$, differently from the kinetic model employed for cellular retinol-binding protein [9].

In this line shape analysis conformational exchange in the apo form, as identified by lower intensity resonances, was considered as a slow outside exchange with equilibrium constant K_{ex} (K_{ex}<1) as described previously[10]. This treatment is equivalent to adding an extra reaction component in slow exchange except that no chemical shift or reaction rate must be assumed for the additional reaction partner[10].This approach should be valid in all cases where the apo resonance is not exchange broadened.

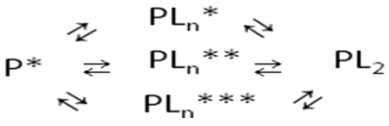
For all the analysed residues kinetic parameters for a two-state binding model show reasonable agreement with experimental line shapes indicative of a binding process involving an initial fast exchange process with an intermediate of relatively low population followed by a slower step.

The first step, initiated by ligand addition, represents a lower affinity interaction with the first ligand molecule, followed by a high affinity binding of the second ligand, consistent with an allosteric control where the second binding event is favored by the binding of the first ligand. This could indicate that the system is endowed with some positive cooperativity, in agreement with previous results[17].

Among residues sensing the two-step binding mechanism, a subgroup (see group A1, Table 1) exhibits measurable chemical shift changes in the first titration steps on the side of the apo-protein, typical of a fast exchange regime ($k_{off,1} \geq 2000\ s^{-1}$), followed by a slow second step ($k_{off,2} \approx 10\text{-}20\ s^{-1}$). The behaviour of D69 is shown as an example in Figure 3a,c. These k_{off} rates point to the presence of a transient intermediate. An evaluation of the population of each species at the different protein:ligand ratios, was obtained from line shape simulation, as reported in Figure 3e. The derived intermediate population reaches 20% at its maximum, and decays to zero once the two-fold stoichiometric amount of ligand has been reached. The residues displaying this behaviour are evenly distributed on the protein, consistent with the fact that they are part of an allosteric system where binding and protein conformational changes are tightly linked

Another subgroup of the residues fitting a two-step binding mechanism (see group A2, Table 1) is characterized by a fast to intermediate $k_{off,1}$ rate ($k_{off,1}$ ~ 400-600 s^{-1}) for the first step, followed by a slow second step ($k_{off,2} \approx 10\text{-}20\ s^{-1}$) (Figure 3b,d). The derived population of the intermediate state appears to be maximal (up to 20%) around a 1:1 P:L ratio and decays to zero once the two-fold stoichiometric amount of ligand has been reached (Figure 3f). Interestingly, mass spectrometry data recorded for different protein:ligand ratios (P:L 1:1, 1:2, 1:5) showed the presence of a non-negligible population of a protein bound to a single bile acid[17], thus suggesting that an intermediate could be the singly ligated species PL_1. As all the residues characterised by the A2 type behaviour (Table 1) are localized in the portal region of the cavity (Figure 4), we suggest that whenever a singly ligated complex is formed, bile acid binds to the superficial site. We do not have evidence whether this equilibrium is part of a productive binding mechanism or represents an off-pathway intermediate. However, the identification of a singly ligated species, where the ligand occupies preferentially the superficial site, is consistent with the finding that in zebrafish liver BABP, the only reported protein of the family exhibiting a 1:1 stoichiometry, the ligand binds at a superficial site, located in the portal region[25].

A large number of residues (see group B, Table 1) is characterized by complex lineshapes showing many shoulders (Figure 5a). These residues are distributed throughout the β-barrel region of the protein structure (Figure 5b). Although the complexity of such line shapes is beyond the scope of simulations, the observation of ligand addition causing shoulders making the apo signal gradually disappear has previously been associated with series of slowly exchanging conformers as described by Scheme 1:



Scheme 1. Suggested parallel binding mechanisms derived by complex behaviour displayed by line shapes.

Here the heterogeneity in the intermediate state (PL_n with n=1,2) can only arise from different apo states which all bind the ligand. In particular this holds for all those cases where the signal with shoulders is also moving upon ligand addition (Figure 3a). Shifting signals associated with pseudo slow-exchange line shapes is also a clear indication of a low population intermediate[10].

Binding heterogeneity is also reflected in the ^{1}H-^{15}N HSQC spectra of enriched GCDA bound to unlabelled protein at different P:L ratios, where multiple resonances are detected for the ligand bound at both sites[17].

The cross sections of a few residues (Table 1) show a holo resonance displaying lower intensity compared to the apo signal (Figure 5c), suggesting the presence of a slow to intermediate exchange between different holo species. Most of the residues displaying this behaviour cluster nicely in the two helices and in A, B, C and J strands (Figure 5d). This pattern points to modulation in protein dynamics after ligand binding consistent with the onset of the equilibrium $PL_1L_2 \rightleftarrows PL_1L_2^*$. As previously described for the conformational equilibria of the apo form, the onset of a similar equilibrium for the holo protein was taken into account in the simulation by using an external K_{ex} value >1 for those residues where the signal intensity was lower in the holo form. This final protein rearrangement, relevant for binding affinity, may occur on a very slow timescale, as previously suggested for human ileal BABP [14]. Interestingly, for one residue (V116) it was possible to detect two holo forms exhibiting a measurable chemical shift difference of 20 Hz between the two holo resonances, corresponding to an upper limit of 125 s^{-1} for the interconversion between PL_1L_2 and $PL_1L_2^*$ (Figure S1). Relaxation dispersion experiments, performed on the holo protein, did not show dispersive behaviour for any residue, thus indicating that conformational rearrangement should indeed occur on a slow time scale, in the order of seconds.

The same line shape analysis was performed, for comparison, on L-BABP (devoid of disulphide bridge), where 42 residues could be analysed, yielding results in complete agreement with the mechanism proposed for L-BABP/S-S.

In L-BABP an intermediate exchange regime characterizes the interconversion between apo conformers, as revealed by the

presence of a high number of residues displaying signals with lower intensity (see Table 2). The simulation of residues exhibiting an apparent slow exchange regime could not be performed employing a single step binding mechanism model, thus suggesting the presence of an intermediate. The line shapes of many residues (see group A, Table 2) located in helical regions and in the barrel, have been fitted with a two-step binding mechanism model, as observed for L-BABP/S-S. A subgroup of residues (see group A2, Table 2) is characterized by a slow k_{off} rate for both steps ($k_{off,1}$ in the range of 20-60 s^{-1} and $k_{off,2}$ in the range of 10-20 s^{-1}) while another subgroup (see group A1, Table 2) by a first step with a fast or a fast to intermediate k_{off} rate ($k_{off,1}$ ~ 100-2000 s^{-1}), followed by a slow second step ($k_{off,2}$ ≈ 10-20 s^{-1}).

In general L-BABP exhibits lower $k_{off,1}$ rates compared to L-BABP/S-S suggesting a higher affinity of L-BABP for GCDA. As an example line shapes simulations of residue S122 are reported in Figure 6 a, b. Many residues, evenly distributed in the protein scaffold (Figure S2), exhibit lineshapes intensities very similar to those of the final holo state around 1:1.6 P:L ratio (Figure 6b). This observation is consistent with an allosteric behaviour and correlates with the high binding cooperativity reported for this protein[12].

The intermediate species formed along the binding pathway in L-BABP are, most likely, interconverting doubly ligated forms, as confirmed by mass spectrometry which does not detect a singly ligated species at any P:L ratio under 1:2. ^{1}H-^{15}N HSQC spectra of the protein recorded during the titration with the ligand support the presence of these intermediates. In fact in some cases the cross peaks describe a curve from apo to holo state for increasing ligand concentrations instead of moving along a straight line between the apo and holo cross peaks (Figure 7). This behaviour is only possible if low-population intermediates with different chemical shifts are present.

Residues characterized by complex lineshapes, showing small shoulders along the titration pathway were additionally detected (see group B, Table 2), again suggesting the presence of slow binding involving intermediates. Finally, residues were detected exhibiting holo resonances with lower intensity with respect to the corresponding apo (Table 2), consistent with the slow (of the order of 110 s^{-1}) rearrangement $PL_1L_2 \rightleftarrows PL_1L_2^*$.

The presented analysis allows to rationalise the behaviours observed for different residues in the two proteins (L-BABP/S-S and L-BABP) contributing to the definition of the following simplified multi-step binding mechanism (Scheme 2):

$$P \rightleftarrows P^* + L \rightleftarrows PL_n + L \rightleftarrows PL_1L_2 \rightleftarrows PL_1L_2^* \quad \text{with n=1,2}$$

Scheme 2 General multistep binding mechanism proposed for bile salts binding to liver BABPs

These results additionally suggest that the presence/absence of a disulphide bridge in L-BABPs is not affecting the general binding mechanism, rather it seems to have an influence on the kinetics of the first exchange process, which in turn translates, in the case of cL-BABP/S-S, into a less cooperative but site-selective system towards glycochenodeoxycholic and glycocholic salts[17]. cL-BABP does not display any site-selectivity, but is an extremely cooperative system[12]. Thus the different behaviour of the two proteins may be related to the bile-acid binding capability, which seems to be directly correlated with cooperativity.

In intracellular proteins, disulfide bonds are generally transiently formed, owing to the reductive nature of the cellular environment. It's plausible that equilibrium between oxidized and reduced forms of L-BABPs, may respond to different needs for ligand uptake.

In conclusion, the combination of dynamic data derived for a lipid carrier from line shape analysis and relaxation dispersion has facilitated the distinction between two fundamentally different mechanisms in response to bile acid binding, namely, induced fit versus conformational selection. The data presented here unambiguously support the conformational selection model for liver BABPs, in which the apo protein samples both the non-competent and binding competent conformations, involving opening at the portal area. Despite the complexity exhibited by a ternary complex, the mechanism involved in the binding of the two ligands could be tackled providing an estimate of the kinetics involved. These findings, while providing a general approach to the analysis of complex binding mechanisms, represent the first step towards the global description of the mechanism of bile acid binding and release within the cell.

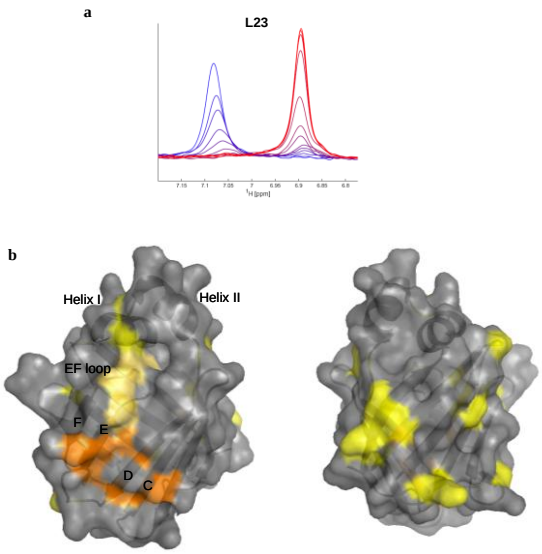


Figure 1. L-BABP/S-S residues showing a low intensity signal for the apo form. a) ^{1}H cross sections through ^{1}H-^{15}N HSQC titration spectra of residue L23. Experimental lines are indicated as solid lines graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration. b) Residues showing low intensity signal for the apo form are mapped in colour on the structure. In the left panel, residues defining a ridge are colour coded from helical region (yellow), through the EF loop (light orange) to the C, D, E strands (orange). The right panel represents the complementary view obtained after a rotation of 180° around z-axis. No specific pattern is here identified and all residues are coloured in yellow. Selected secondary structural elements are labelled.

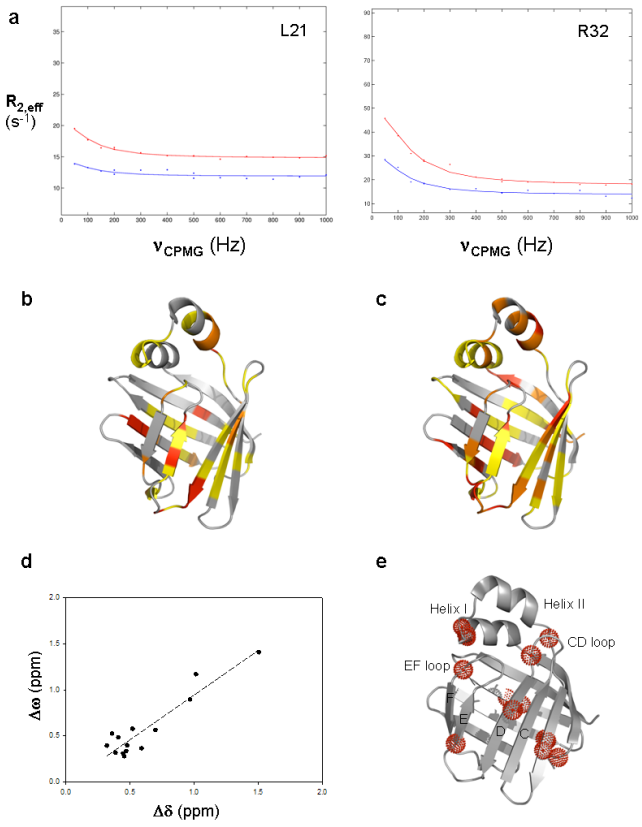


Figure 2. Comparison of chemical shift differences deduced from relaxation dispersion and titration data. a) Representative ^{15}N relaxation dispersion curves obtained for L21 and R32 at 600 MHz (blu) and 900 MHz (red). b) Residues showing exchange contribution to backbone ^{15}N relaxation dispersion are mapped on the structure and are coloured according to their $\Delta\omega$ values using a scale from yellow (low values, <0.8 ppm) to red (high values, >1.5 ppm). c) ^{15}N Chemical shift differences between apo and holo proteins are mapped on the structure and are coloured according to their $\Delta\delta$ values using a scale from yellow (low values, <0.8 ppm) to red (high values, >1.5 ppm). Data are shown only for values of $\Delta\delta$ >0.3 ppm, which is the smallest value detected from relaxation dispersion analysis. d) Correlation between $\Delta\omega$ values, obtained from relaxation dispersion measurements, and differences ($\Delta\delta$) between ^{15}N chemical shifts available for apo and holo protein states observed for residues: S3, A20, L21, A22, I34, I40, Q41, Q56, T59, A68, K76, V102, E109, T110 (slope 0.97, R^2 0.98). e) Residues for which the linear correlation was obtained are indicated as red dotted spheres on the structure. Selected structural elements are labelled

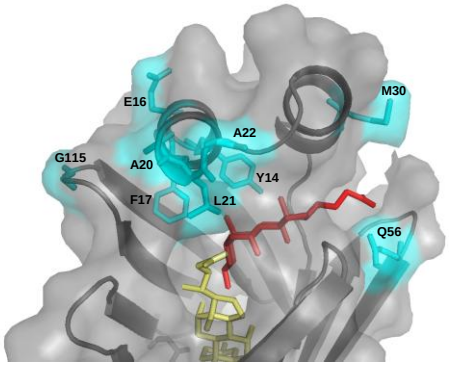


Figure 4. Mapping of residues characterized by a fast to intermediate k_{off} rate in the first binding step (eq. [1]). Residues Y14, E16, F17, A20, L21, M30, Q56, G115 are coloured in cyano and labelled on the surface of L-BABP/S-S in complex with two bile salts. The superficial ligand is coloured in red.

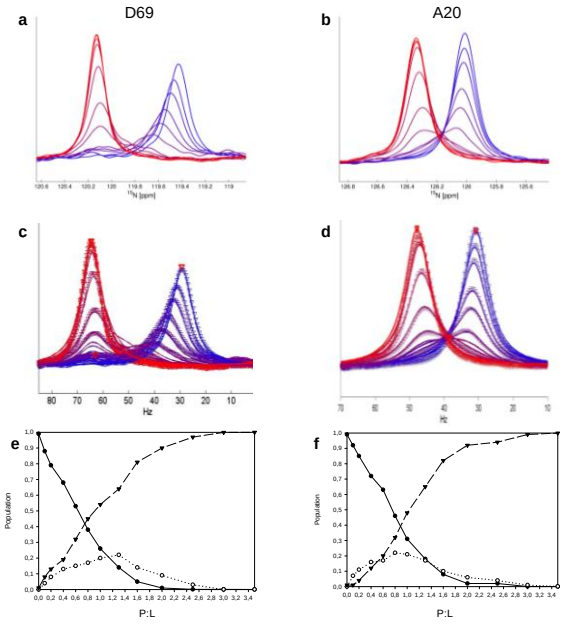


Figure 3 Line shape simulation of L-BABP/S-S titrated with GCDA using a two step binding mechanism (eq.[1]). a) ^{15}N cross sections through ^{1}H-^{15}N HSQC titration spectra of residue D69. b) ^{15}N cross sections through ^{1}H-^{15}N HSQC titration spectra of residue A20. Lines are graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration c,d) Line shape simulation performed with NMRKIN for residue D69 (c) and A20 (d) (triangles). Calculated off-rates : $k_{off,1}$ = 3000 s^{-1}, $k_{off,2}$ = 10 s^{-1} for D69; $k_{off,1}$ = 500 s^{-1} and $k_{off,2}$,= 30 s^{-1} for A20. e,f) Populations of species P (filled circle), I (empty circle) and PL_2 (triangle), as deduced from the line-shape analysis, are reported for residue D69 (e) and A20 (f).

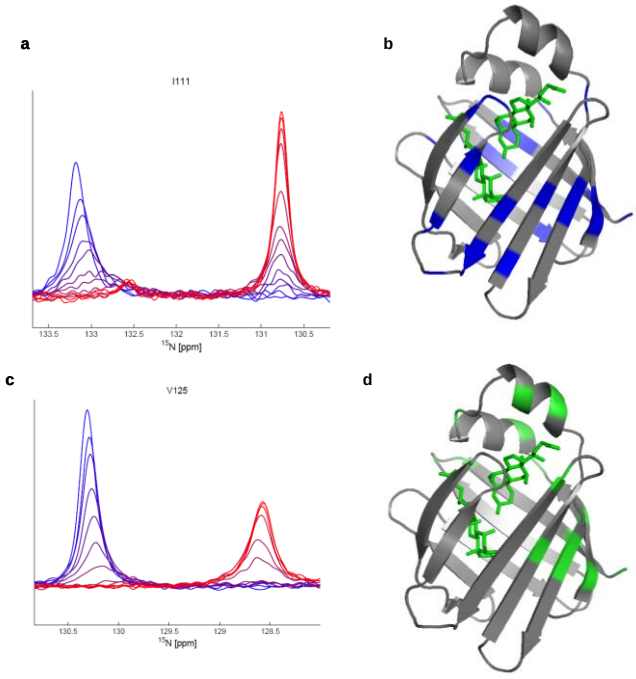


Figure 5. Residues characterized by conformational averaging. a) I111 ^{15}N cross section exhibiting several shoulders, indicated by arrows, in intermediate titration steps. Lineshapes are graded from blue (apo protein) to red (holo-protein) for subsequent steps of the titration. b) Residues showing lineshapes with shoulders after the addition of the ligand (T5, W6, Y9, E12, N13, I34, V38, E39, F47, T50, N60, F62, T63, G65, A68, T71, D74, G75, A85, F96, V102, T110, I111, L118, I119, K123, V125) are labelled in blue on the holo L-BABP/S-S complex. c) V125 ^{15}N cross sections exhibiting low intensity signal for the holo form. Lineshapes are graded from blue (apo protein) to red (ligand-bound protein) for subsequent steps of the titration. d) Residues exhibiting low intensity signal in the holo form. (T5, V8, E12, E15, A22, L27, M30, V38, E39, T50, Q56, G114, V116, S122, V125) are labelled in green on the holo L-BABP/S-S complex.

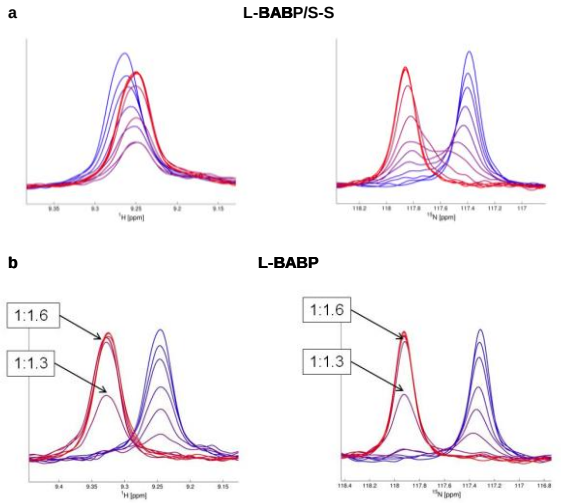


Figure 6. ^{1}H and ^{15}N cross sections through S122 cross peak for subsequent steps of the titration of L-BABP/S-S (upper panel) and L-BABP (lower panel) with GCDA. Experimental line-shapes are indicated as solid lines graded from blue (apo) to red (holo protein) for subsequent steps of the titration. a) ^{1}H and ^{15}N cross sections through residue S122 in L-BABP/S-S protein. Calculated off-rates: $k_{off,1}$ = 3000 s^{-1} and $k_{off,2}$ = 15 s^{-1}. b) ^{1}H and ^{15}N cross sections through residue S122 in L-BABP. Calculated off-rates: $k_{off,1}$ = 100 s^{-1} and $k_{off,2}$ = 10 s^{-1}. In b) the lineshapes corresponding to P:L ratios of 1:1.3 and 1:1.6 are labelled.

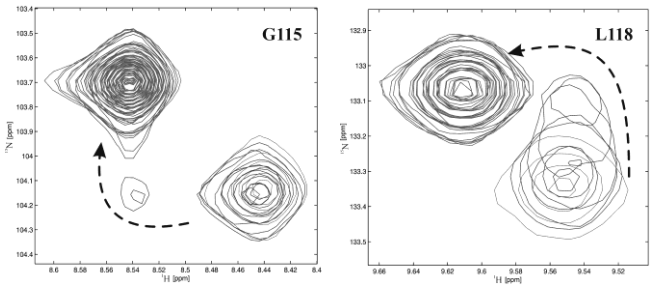


Figure 7. Superposition of the thirteen ^{1}H ^{15}N HSQC spectra, recorded for subsequent steps of L-BABP titration with GCDA, showing cross-peaks of residue G115 a) and L118 b) moving on a 'curve'. The dotted arrow indicates the path followed by the pick along the titration going from apo to holo protein.

Table 1 Grouping of residues of L-BABP/S-S according to binding mechanism they sense as suggested by lineshapes analysis

Kinetic model	residues
K_{ex} on Apo	Y9, L23, I34, F47, V48, N60, F62, G65, A68, D69, T71, D74, G75, A85, F96, V102, K103, T110, I111, G115, L118, K123
Two step binding mechanism (**A**)	
A1. Fast/slow exchange regime	S3, V8, E15, A22, L23, L27, V48, D69, G87, K103, G114, V116, S122
A2. Fast to intermediate / slow exchange regime	Y14, E16, F17, A20, L21, M30, Q56, G115
Parallel binding mechanism (**B**)	T5, W6, Y9, E12, I34, V38, E39, F47, T50, N60, F62, T63, G65, A68, T71, D74, G75, A85, F96, V102, T110, I111,L118, I119, K123, V125
K_{ex} on Holo	T5, V8, E12, E15, A22, L27, M30, V38, E39, T50, Q56, G114, V116, S122, V125

Table 2 Grouping of residues of L-BABP according to binding mechanism they sense as suggested by lineshapes analysis

Kinetic model	residues
K_{ex} on Apo	Q7, N13, F17, L21, T63, G65, A68, G75, K79, A85, L89, T112, G115, V116, L118, I119
Two step binding mechanism (**A**)	
A1. Fast-fast to intermediate/slow exchange regime	G4, L23, E25, L27, M30, S61, K103, L118
A2. Slow/slow exchange regime	Q7, V8, N13, E15, F17, A20, L21, V38, Q56, T112, G114, G115, V116, S122
Parallel binding mechanism (**B**)	W6, E39, I40, F47, T50, T63, G65, A68, G75, K79, A85, T112, I119, V125
K_{ex} on Holo	V8, A20, L27, M30, K35, V38, I40, F47, S51, Q56, V116, V125

Experimental Section

Sample preparation. The proteins were expressed as previously reported [8, 17]. All the protein preparations were checked by 1D ^{1}H NMR prior to and after delipidation, and reproducible spectra were always obtained for the apo and undelipidated forms. The molecular weight of the two proteins and the extent of labelling was verified by MALDI mass analysis. NMR samples contained 0.5 mM ^{15}N-BABP protein dissolved in 30 mM phosphate (Na_2HPO_4/NaH_2PO_4), 90%H_2O-10%D_2O. The pH of the solutions was 7.2. Protein stock solution concentrations were determined by UV. Bile salts were purchased from Sigma. The concentration of the ligand stock solutions were determined by measuring dry weights using a microbalance. Ligands were dissolved in the same 30 mM phosphate buffer (pH 7.2). For protein-ligand titration experiments, glycochenodeoxycolate was added to the protein in small amounts to achieve saturation within 10 steps (P:L ratios employed 1:0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.3, 1.6, 2.0), and three additional aliquots were added to ensure the end of titration (P:L ratios 1:2.5, 3.0, 3.5).

NMR assignments. NMR spectra were acquired at 298 K on Bruker spectrometers DMX 500 and Avance III 600 equipped with a 5mm TCI cryoprobe and Z-field gradient. Data were processed with NMRPipe[26] and visualized by NMRView[27]. For resonance assignment of L-BABP/S-S complexed with GCDA ^{1}H-^{15}N 3D TOCSY and NOESY-HSQC were collected. Standard pulse sequence schemes using pulsed field gradients were employed to achieve suppression of the solvent signal and spectral artefacts. Proton chemical shifts were referenced to external 3-(trimethylsilyl)-3,3,2,2-d4-propionic acid, while nitrogen chemical shifts were referenced indirectly as described[28].

^{1}H and ^{15}N protein resonance assignment has been reported in Table S1 in Supporting information.

Line shape analysis. For lineshape analysis spectra were recorded on a Bruker AVANCE 500 MHz spectrometer equipped with a cryogenically cooled 5mm TXI probe. The operating field was chosen to keep exchange in a faster exchange regime than it would otherwise be at higher field strengths. After each ligand addition, a high-resolution sensitivity enhanced $^{1}H/^{15}N$-HSQC spectrum was recorded with 2048 × 800 data points. Spectra were processed with zero-filling and sine window functions.

For residues which were not subject to chemical shift degeneracy, cross sections in both spectral dimensions were extracted from all spectra by carefully selecting slices containing the signal maxima. NMRLab software [29] was employed for spectra processing. Line shapes for both dimensions were simulated simultaneously for selected kinetic models using the set of thirteen experimental spectra. NMRKIN was used to calculate the time domain signal for sections of HSQC spectra for a given kinetic mechanism assuming steady-state line shapes as described by the equations of Gutowsky and McConnell [30, 31]. Chemical shifts of signals, line widths, populations, and kinetic rates were adjusted iteratively to achieve consistency with experimental data. In the slow exchange regime, only an upper limit of the off-rate can be assigned because the line shapes depend only slightly on the rate. Residues were assigned to the model that best described their line shapes.

Kinetic model implemented in the NMRKIN software for Lineshape Simulations. Lineshapes were simulated as described in the original NMRKIN publication [10]. For the case of a two step binding mechanism an additional model was implemented which is described in the following. As described in [10] the total magnetization is obtained by adding up the individual components:

$$M_{tot} = I^T M = iC1^T A^{-1} P$$

with $A = 2\pi i\,(Iv - W_o) + R_2 + K$ where I is the unity matrix, v is a variable frequency defining the frequency range, W_o is a diagonal matrix with Larmor frequencies ω, P a column vector with the populations of the individual states, R_2 is a diagonal matrix with transverse relaxation rates R = $1/T_2^* = \pi$ LW (line width at half height), K is a matrix containing the rate constants (rate matrix).
NMRKIN calculates lineshapes in the time domain according to

$$F(t) = I \exp\,[(-K + 2\pi i W_o - R_2)\,t]\,P$$

followed by the Fourier transformation to obtain the spectrum.
The mechanism is defined by the matrix K and by associations between rates and populations. For the simulations in this work we implemented a mechanism for two-stages binding as described by

$$P + L \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} PL_n + L \underset{k_{32}}{\overset{k_{23}}{\rightleftarrows}} PL_2L_2,$$

where ligand is bound in two subsequent steps.
We assume that there is no direct conversion of free protein P into the product with two ligands PL_2L_2, i.e. the rates k_{13} and k_{31} have a value of zero.
The rate matrix for this case is:

$$K = \begin{bmatrix} k'_{12} + k_{13} & -k_{21} & -k_{31} \\ -k'_{12} & k_{21} + k'_{23} & -k_{32} \\ -k_{13} & -k'_{23} & k_{31} + k_{32} \end{bmatrix} \equiv \begin{bmatrix} k'_{12} & -k_{21} & 0 \\ -k'_{12} & k_{21} + k'_{23} & -k_{32} \\ 0 & -k'_{23} & k_{32} \end{bmatrix}$$

Here $k'_{12} = k_{12}[L] = \frac{1}{\tau_{12}}$ and $k'_{23} = k_{23}[L_2L_2] = \frac{1}{\tau_{23}}$ depend on the concentration of the concentration of ligand added.
For equilibrium conditions we get:

$k_{12}[L]p(P) = k_{21}p(PL_n)$ and therefore

$$\tau_{12} = \tau_{21}\frac{p(P)}{p(PL_n)} \quad (1)$$

and the same for the second step of the mechanism:

$k_{23}[L]p(PL_n) = k_{32}p(PL_2L_2)$ and therefore

$$\tau_{23} = \tau_{32}\frac{p(PL_n)}{p(PL_2L_2)} \quad (2)$$

After eliminating $p(PL_n)$ from equation (1) and (2) we get

$$\tau_{21}\tau_{32}\frac{p(P)}{p(PL_2L_2)} = \tau_{23}\tau_{12}$$

with $p(P) + p(PL_n) + p(PL_2L_2) = 1$

$$\tau_{23} = \tau_{32}\frac{1 - p(PL_2L_2) - p(P)}{p(PL_2L_2)}$$

The NMRKIN mechanism file was set up with these simplifications.

Relaxation Dispersion. Experiments were acquired on Varian Inova spectrometers equipped with room temperature probe heads at two static magnetic fields, corresponding to proton Larmor frequencies of 600 and 900 MHz using a relaxation compensated CPMG dispersion experiment performed in a constant time manner [32]. Spectra were collected as a series of twelve two-dimensional spectra with CPMG field strengths ν_{CPMG}: 50, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900 and 1000 Hz with repeat experiments recorded at fields of 200 and 500 Hz. Eight scans were recorded per FID with a 3 s recycle time. The constant time period was set to 40 ms; the reference experiment was carried out by omitting the CPMG period. Spectra were acquired with at least 1024/2=512 complex points for 600 MHz data and 1210/2=605 complex points for 900 MHz data, in F2 and F1 dimensions, respectively. Processing and analysis of the NMR spectra were performed employing NMRlab software [29].
For varying echo times, relaxation dispersion curves are obtained which depend on the chemical shift differences ($\Delta\omega$), the population of the two states, and the exchange rate k_{ex}. Exchange rates (k_{ex}) were determined by fitting relaxation dispersion curves to the general expression for the phenomenological transverse relaxation rate constant in the case of two-site exchange[19] .
Pymol (DeLano, W.L. The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA, USA) was used for graphical representation of the results employing 1tvq and 2jn3 structures.

Acknowledgements

Authors are indebted to Lucia Zetta for useful discussions. This research was possible thank to the University of Verona for financial support in the acquisition of the NMR Bruker Avance 600 MHz spectrometer and to Cariverona Foundation for the acquisition of the cryoprobe. We acknowledge access to instrumentation at HWB-NMR (University of Birmingham) supported by EU-NMR (RII3-026145). L.R. thanks CNR-RSTL 2007 (Code N° 779) for financial support. CC was supported by a short term fellowship from EMBO (ref: ASTF 41-2009) and by CNR short term mobility 2008.

Keywords: intracellular lipid binding proteins, NMR, line shape analysis, molecular recognition, bile acids

[1] S. M. Houten, M. Watanabe, J. Auwerx, *Embo J* **2006**, *25*(7), 1419-1425.
[2] C. Thomas, R. Pellicciari, M. Pruzanski, J. Auwerx, K. Schoonjans, *Nat Rev Drug Discov* **2008**, *7*(8), 678-693.
[3] G. V. Richieri, R. T. Ogata, A. W. Zimmerman, J. H. Veerkamp, A. M. Kleinfeld, *Biochemistry* **2000**, *39*(24), 7197-7204.
[4] T. Hanhoff, C. Lucke, F. Spener, *Mol Cell Biochem* **2002**, *239*(1-2), 45-54.
[5] A. Chmurzynska, *J Appl Genet* **2006**, *47*(1), 39-48.
[6] O. Toke, J. D. Monsey, G. T. DeKoster, G. P. Tochtrop, C. Tang, D. P. Cistola, *Biochemistry* **2006**, *45*(3), 727-737.
[7] G. P. Tochtrop, G. T. DeKoster, D. F. Covey, D. P. Cistola, *J Am Chem Soc* **2004**, *126*(35), 11024-11029.
[8] L. Ragona, M. Catalano, M. Luppi, D. Cicero, T. Eliseo, J. Foote, F. Fogolari, L. Zetta, H. Molinari, *J Biol Chem* **2006**, *281*(14), 9697-9709.
[9] T. Mittag, L. Franzoni, D. Cavazzini, B. Schaffhausen, G. L. Rossi, U. L. Gunther, *J Am Chem Soc* **2006**, *128*(30), 9844-9848.
[10] U. L. Gunther, B. Schaffhausen, *J Biomol NMR* **2002**, *22*(3), 201-209.
[11] G. P. Tochtrop, K. Richter, C. Tang, J. J. Toner, D. F. Covey, D. P. Cistola, *Proc Natl Acad Sci U S A* **2002**, *99*(4), 1847-1852.
[12] M. Pedo, M. D'Onofrio, P. Ferranti, H. Molinari, M. Assfalg, *Proteins* **2009**, *77*(3), 718-731.
[13] A. M. Rea, V. Thurston, M. S. Searle, *Biochemistry* **2009**.
[14] O. Toke, J. D. Monsey, D. P. Cistola, *Biochemistry* **2007**, *46*(18), 5427-5436.
[15] T. Eliseo, L. Ragona, M. Catalano, M. Assfalg, M. Paci, L. Zetta, H. Molinari, D. O. Cicero, *Biochemistry* **2007**, *46*(44), 12557-12567.
[16] G. P. Tochtrop, J. L. Bruns, C. Tang, D. F. Covey, D. P. Cistola, *Biochemistry* **2003**, *42*(40), 11561-11567.
[17] C. Cogliati, S. Tomaselli, M. Assfalg, M. Pedo, P. Ferranti, L. Zetta, H. Molinari, L. Ragona, *Febs J* **2009**, *276*(20), 6011-6023.
[18] T. Mittag, B. Schaffhausen, U. L. Gunther, *J Am Chem Soc* **2004**, *126*(29), 9017-9023.
[19] T. Mittag, B. Schaffhausen, U. L. Gunther, *Biochemistry* **2003**, *42*(38), 11128-11136.
[20] M. D'Onofrio, L. Ragona, D. Fessas, M. Signorelli, R. Ugolini, M. Pedo, M. Assfalg, H. Molinari, *Arch Biochem Biophys* **2009**, *481*(1), 21-29.
[21] T. Mittag, S. Orlicky, W. Y. Choy, X. Tang, H. Lin, F. Sicheri, L. E. Kay, M. Tyers, J. D. Forman-Kay, *Proc Natl Acad Sci U S A* **2008**, *105*(46), 17772-17777.

[22] D. D. Boehr, D. McElheny, H. J. Dyson, P. E. Wright, *Science* **2006**, *313*(5793), 1638-1642.
[23] U. Brath, M. Akke, *J Mol Biol* **2009**, *387*(1), 233-244.
[24] A. E. Jenkins, J. A. Hockenberry, T. Nguyen, D. A. Bernlohr, *Biochemistry* **2002**, *41*(6), 2022-2027.
[25] S. Capaldi, M. Guariento, G. Saccomani, D. Fessas, M. Perduca, H. L. Monaco, *J Biol Chem* **2007**, *282*(42), 31008-31018.
[26] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J Biomol NMR* **1995**, *6*(3), 277-293.
[27] B. A. Johnson, *Methods Mol Biol* **2004**, *278*, 313-352.
[28] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, B. D. Sykes, *J Biomol NMR* **1995**, *6*(2), 135-140.
[29] U. L. Gunther, C. Ludwig, H. Ruterjans, *J Magn Reson* **2000**, *145*(2), 201-208.
[30] H. S. Gutowsky, D. McCall, C. Slichter, *The Journal of chemical physics* **1953**, *21*, 279-292.
[31] H. M. McConnell, *J. Chem. Physics* **1958**, *28*, 430-431.
[32] A. G. Palmer, 3rd, C. D. Kroenke, J. P. Loria, *Methods Enzymol* **2001**, *339*, 204-238.

Received: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))

8.Dynamics and interaction: Additional results

8.1 Lineshape simulation: from “one” to “two binding step” mechanisms

In the following paragraph, as a compendium of the previous paper, a short example is given of the procedure employed for the selection of the “appropriate” binding model for data analysis.

Mechanism files used in NMRKIN, with their respective abbreviation and meaning, are listed in Table 8.1.

m11.m	$A \rightleftharpoons B$
m21.m	$P + L \rightleftharpoons PL$
m2121.m	$P + L \rightleftharpoons PL_n + L \rightleftharpoons PL_1L_2$

Table 8.1. Mechanism files used in NMRKIN software.

Figure 8.1 reports the experimental lineshapes of 8 residues (Y14, E16, F17, A20, L21, M30, Q56, G115), obtained during titration of cL-BABP/S-S with GCDA.

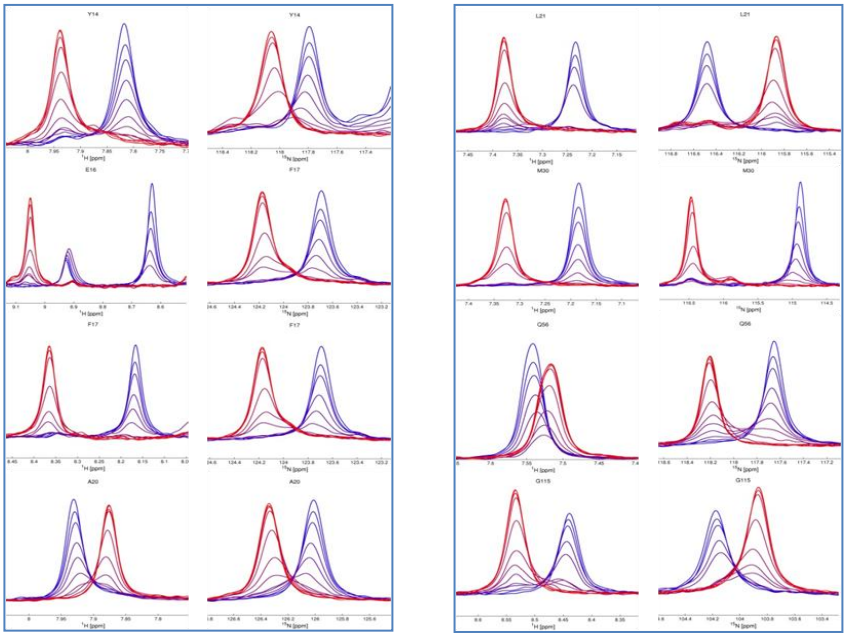


Figure 8.1. ^{1}H and ^{15}N cross sections through ^{1}H-^{15}N HSQC titration spectra of cL-BABP/S-S residues Y14, E16, F17, A20, L21, M30, Q56, G115 firstly simulated with a simple one step model. Experimental lines are indicated as solid lines graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration.

These cross-sections were initially considered as typical slow exchange lineshapes that could be simulated with a second order reaction of the type:

$$P + L \rightleftharpoons PL \quad (1)$$

where P is the apo protein; L is the ligand and PL is the holo protein.

The lineshapes simulation, through the m21 mechanism, did not fit with the experimental lines observed during the titration.

The lineshape analysis, performed for comparison on cL-BABP (devoid of disulphide bridge), showed that also in this case the simulation of experimental lines could not be performed considering a single binding step mechanism. Thus, what was initially considered as a simple slow exchange process, in which the line corresponding to P decreases, while the line corresponding to PL_1L_2 increases, without any shift of the signal and any intermediate forms, did not properly describe the interaction.

The presence of an elusive intermediate had to be taken into account to obtain a good lineshape simulation.

The m2121 kinetic model was subsequently implemented to simulate experimental lineshapes. The model represented a two step binding mechanism, by which cL-BABPs forms a ternary complex with two bile acid molecules:

$$P^* + L \underset{k_{off,1}}{\rightleftharpoons} PL_n + L \underset{k_{off,2}}{\rightleftharpoons} PL_1L_2 \quad (2)$$

where P^* represents the protein in one of its apo conformation; L represents the ligand; PL_n represents a generic intermediate which could derive either from differences in protein conformations or in the type of bound species, and PL_1L_2 represents the ternary complex.

Some examples of simulated lineshapes, not included into the previous paper, are here reported.

In Figure 8.2 the simulated ^1H and ^{15}N lineshapes are reported for residue L21 of cL-BABP/S-S.

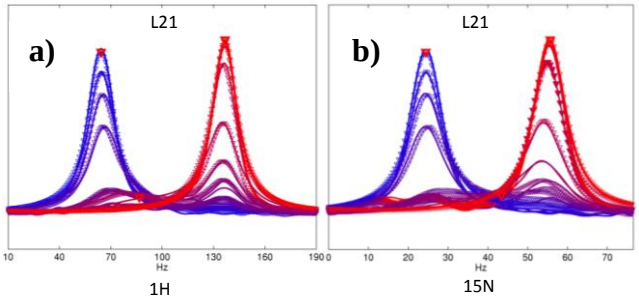


Figure 8.2 Lineshape simulation of L21 residue, obtained through ^1H-^{15}N HSQC titration of cL-BABP/S-S with GCDA, using a two step binding mechanism (eq.(2)). ^1H and ^{15}N cross sections (panel **a** and **b**, respectively), are indicated with simple lines. Simulated lineshapes are indicated with triangles. Lines are graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration. Calculated off-rates are reported in Table 2

Figure 8.3 reports examples of ^{15}N lineshape simulations of cL-BABP/S-S residues characterised by $k_{off,1}$ values typical of a fast exchange regime ($k_{off,1} \geq 2000$ s^{-1}) and $k_{off,2}$ values characteristic of a slow exchange regime (~ 15-20 s^{-1}).

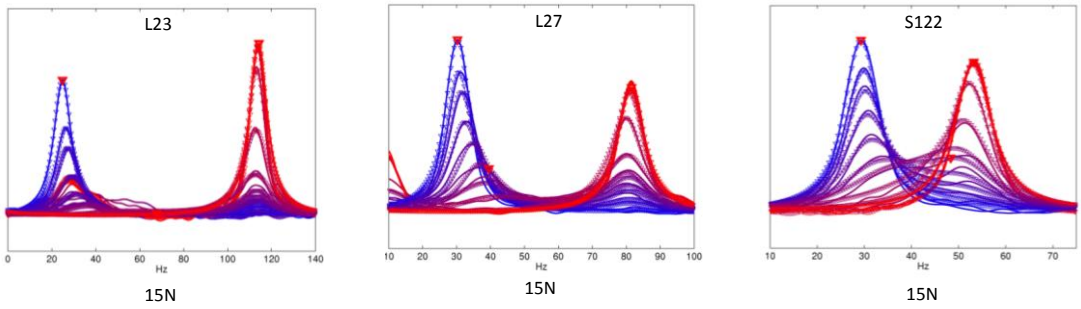


Figure 8.3 Lineshape simulation of L23, L27 and S122 residues, obtained through ^{1}H-^{15}N HSQC titration of cL-BABP/S-S with GCDA, using a two step binding mechanism (eq.(2)). ^{15}N cross sections are indicated with simple lines. Simulated lineshapes are indicated with triangles. Lines are graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration.

The kinetic constants, obtained for the residues of Figure 8.3 are reported in Table 8.1.

cL-BABP/S-S	$k_{off,1}(s^{-1})$	$k_{off,2}(s^{-1})$	$K_{ex}(s^{-1})$
L23	2500	20	0,89
L27	2000	20	1,14
S122	3000	15	1,02

Table 8.1 Examples of cL-BABP/S-S residues with a fast $k_{off,1}$ rate, a slow $k_{off,2}$ rate and various equilibrium constants depending on a pre-existing equilibrium among apo forms (K_{ex}<1) or on an equilibrium among holo forms (K_{ex}>1).

Figure 8.4 reports examples of ^{15}N lineshape simulations of cL-BABP/S-S residues characterized by $k_{off,1}$ values typical of a fast to intermediate exchange regime ($k_{off,1}$ ~ 300-600 s^{-1}) and a $k_{off,2}$ values characteristic of a slow exchange regime (~ 10-20 s^{-1}).

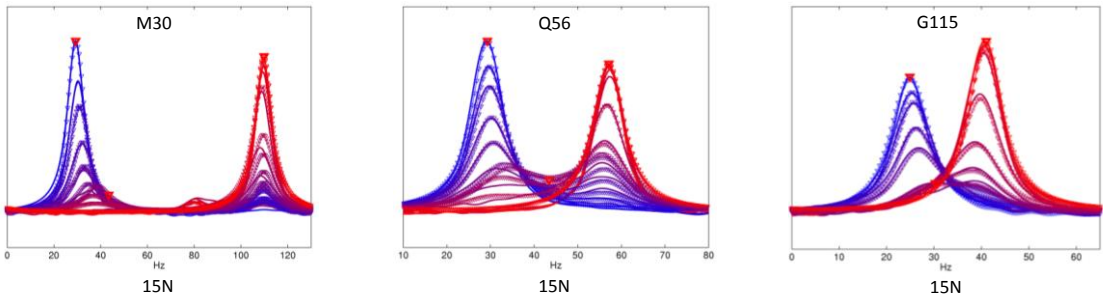


Figure 8.4 Lineshape simulation of M30, Q56 and G115 residues, obtained through ^{1}H-^{15}N HSQC titration of cL-BABP/S-S with GCDA, using a two step binding mechanism (eq.(2)). ^{15}N cross sections are indicated with simple lines. Simulated lineshapes are indicated with triangles. Lines are graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration.

The kinetic constants, obtained for the residues of Figure 8.4 are reported in Table 8.2.

cL-BABP/S-S	$k_{off,1}(s^{-1})$	$k_{off,2}(s^{-1})$	$k_{ex}(s^{-1})$
L21	300	15	1
M30	400	10	1,05
Q56	600	20	1,08
G115	500	20	0,89

Table 8.2 Examples of cL-BABP/S-S residues with a fast to intermediate $k_{off,1}$ rate, a slow $k_{off,2}$ rate and various equilibrium constants depending on a pre-existing equilibrium among apo forms (K_{ex}<1) or on an equilibrium among holo forms (K_{ex}>1).

Examples of ^{15}N lineshape simulations obtained for cL-BABP (devoid of disulphide bridge) are reported in Figure 8.5 for residues of characterised by $k_{off,1}$ values typical of a fast to intermediate exchange regime ($k_{off,1}$ ~ 300-600 s^{-1}) and $k_{off,2}$ values characteristic of a slow exchange regime (~ 10-20 s^{-1}).

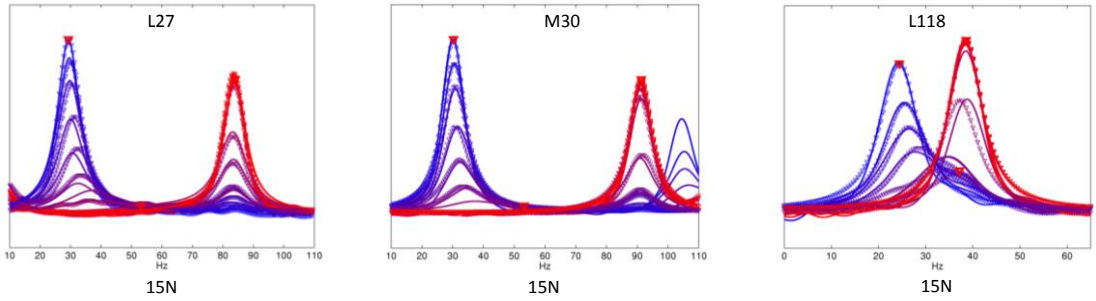


Figure 8.5 Lineshape simulation of L27, M30 and L118 residues, obtained through ^{1}H-^{15}N HSQC titration ofcL-BABP with GCDA, using a two step binding mechanism (eq.(2)). ^{15}N cross sections are indicated with simple lines. Simulated lineshapes are indicated with triangles. Lines are graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration.

The kinetic constants, obtained for the residues of Figure 8.5 are reported in Table 8.3.

cL-BABP	$k_{off,1}(s^{-1})$	$k_{off,2}(s^{-1})$	$K_{ex}(s^{-1})$
L27	100	10	1,15
M30	100	10	1,15
L118	2000	10	0,96

Table 8.3 Examples of cL-BABP residues with a fast to intermediate $k_{off,1}$ rate, a slow $k_{off,2}$ rate and various equilibrium constants depending on a pre-existing equilibrium among apo forms (K_{ex}<1) or on an equilibrium among holo forms (K_{ex}>1).

Figure 8.6 reports examples of ^{15}N lineshape simulations of cL-BABP residues characterised by $k_{off,1}$ ($k_{off,1}$ ~ 20-60 s^{-1}) and $k_{off,2}$ ($10s^{-1}$) values typical of a slow exchange regime.

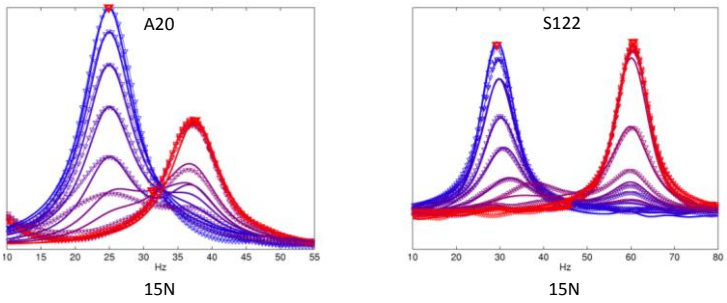


Figure 8.6 Lineshape simulation of A20 and S122 residues, obtained through ^{1}H-^{15}N HSQC titration of cL-BABP with GCDA, using a two step binding mechanism (eq.(2)). ^{15}N cross sections are indicated with simple lines. Simulated lineshapes are indicated with triangles. Lines are graded from blue (apo protein) to red (holo protein) for subsequent steps of the titration.

The kinetic constants, obtained for the residues of Figure 8.6 are reported in Table 8.4.

cL-BABP	$k_{off,1}(s^{-1})$	$k_{off,2}(s^{-1})$	$K_{ex}(s^{-1})$
A20	20	5	1,34
S122	60	15	1,1

Table 8.4 Examples of cL-BABP residues with both slow $k_{off,1}$ and $k_{off,2}$ rates and various equilibrium constants (K_{ex}) depending on a pre-existing equilibrium among apo forms (K_{ex}<1) or on an equilibrium among holo forms (K_{ex}>1).

Table 8.5 lists the MATLAB routines used in NMRKIN software to simulate experimental lineshapes of cL-BABP/S-S and cL-BABP titrated with GCDA.

	Calling routine
mklspar.m	Read and check parameters
testpar.m	Sample parameter file
redat.m	Read experimental data
wdwf.m	Calculate window funcion[a]

[a] Apodization routine from NMRlab

Table 8.5 Routines used in NMRKIN software.

8.2 Relaxation Dispersion experiments at various GCDA concentrations

The CPMG ^{1}H-^{15}N HSQC spectra were recorded at two different fields, 600MHz and 900 MHz, for (i) the apo cL-BABP/S-S protein, (ii) the protein complexed with GCDA at 1:0.1 P:L ratio, (iii) the protein complexed with GCDA at 1:1 P:L ratio and (iv) the protein complexed with GCDA at 1:2.5 P:L ratio. The NMR relaxation dispersion profile of cL-BABP/S-S, at different bile acid concentrations and the comparison among the obtained chemical shifts, give the possibility to provide a k_{ex} (exchange rate) and an hypothetical mechanism of reaction.

The presence of significant μs-ms conformational exchange processes has been observed in early stages of the titration, at 0% and 10% (P:L = 1:0.1) of the ligand. At highest concentrations of GCDA relaxation dispersion was no more observable, possibly because exchange does not fit with the timescale detected by relaxation dispersion experiment.

Relaxation dispersion profiles at the different concentrations of GCDA are reported in Figure 8.7.

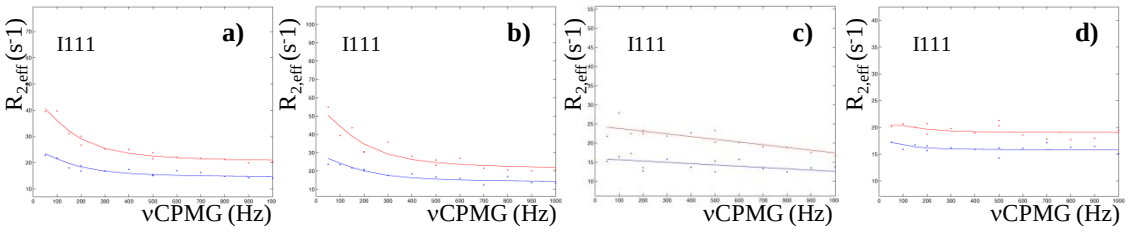


Figure 8.7 NMR CPMG (Carr-Purcell-Meiboon-Gill) Relaxation Dispersion profile. $R_{2,eff}$ (s^{-1}): relaxation rate constant plotted versus ν_{CPMG} (Hz). The data were collected at two different (B_0) magnetic fields: 14,1 Tesla (in blue) and 21,12 Tesla (in red). Panel **a** shows the relaxation dispersion profile of residue I111 in the apo cL-BABP/S-S; panel **b** shows relaxation dispersion profile of residue I111 in the complex cL-BABP/S-S:GCDA at 1:0.1 P:L ratio; panels **c** and **d** show how relaxation dispersion is no more visible for residue I111 in the complex cL-BABP/S-S:GCDA at 1:1 and 1:2.5 P:L ratio.

Tables 8.6-8.9 report k_{ex} rates and $\Delta\omega$ extracted for the residues exhibiting non flat relaxation dispersion profiles. Obtained data are derived from analysis at four different protein:ligand ratios. The data have been analysed, up to now, only with a two-site exchange model which does not reflect the complexity of the involved mechanism.

Further analysis and experiments are done in order to improve the relaxation dispersion analysis.

cL-BABP/S-S apo	$k_{ex}(s^{-1})$	Δω(Hz)	χ^2
E16	733	90.47	2.7
K19	917.8	96.78	3.8
A20	755.9	115.29	3.7
L21	752.3	104.68	2.9
A22	1135.5	197.67	1.8
L23	704.7	100.96	2.9
I28	587.4	121.47	1.8
K29	789.4	137.29	5
M30	484.3	322.6	2.3
R32	650.5	439.45	13.9
D33	745.9	420.59	26.8
I34	158.8	811.61	12.1
K35	774.8	136.76	5.3
V38	194.4	516.04	8.9
I40	116	62.69	6.6
Q41	596	217.04	13.5
F47	970.2	158.97	17.3
V48	729	116.46	6.7
T50	892	435.78	23.1
S51	427.5	540.73	33.9
Q56	683.4	136.28	19.3
T59	37.4	69.67	3.9
S61	501	336.72	4.6
F62	663.5	135.42	12
T63	817	626.34	28
L64	10.6	995.11	18
G65	762.4	167.09	38.5
A68	4.7	628.34	8.4
D69	1074.6	147.8	38.4
T71	20.3	873.76	20.1
G75	534.5	148.36	32.1
K76	579.7	662.81	85.1
H83	1961.2	320.5	27
L84	175.2	440.03	6.8
S97	648.8	300.19	11.5
V102	2.9	568.4	8.2
K103	885.7	234.15	24.3

E106	942.6	64.78	7
V108	585.4	221.42	13.7
E109	843.8	141.2	43
T110	1719.9	181.49	11.3
I111	929.9	149.55	38.1
K123	1033.4	188.91	18.5
V125	3629.5	137.64	5

Table 8.6 k_{ex} and $\Delta\omega$ values obtained for the residues of apo cL-BABP/S-S showing a RD profile. The obtained parameter χ^2, indicative of the goodness of the fit, is reported in the third column.

cL-BABP/S-S GCDA 1/0.1	**$k_{ex}(s^{-1})$**	**$\Delta\omega$(Hz)**	**χ^2**
E16	787.4	136.96	5.2
K19	725	563.65	10.3
A20	616.2	138.37	4.2
L21	382.3	378	2.7
L23	537.1	585.52	8.3
L27	452.4	164.15	4.2
K29	788.4	373.82	3.4
M30	10.4	838.98	3.2
D33	558	175.44	28.4
I34	11.6	819.86	12.5
I37	696.2	681.87	7
V38	63.4	921.77	8.9
I40	470.6	411.79	17.3
F47	63.7	243.72	13.2
V48	792.8	251.83	6.3
V49	674.3	303.24	13.4
Q56	613.3	414.04	15.2
T63	2016.4	389.26	22.1
K66	494.7	538.98	3.6
D69	828.8	558.99	9
K76	110.3	573.69	19.5
L84	756	581.76	12.1
F96	388.7	660.53	21.7
S97	22.2	899.9	27.7
V102	23.4	783.48	24.7
E106	610.8	78.53	1.8
V108	9.3	292.73	9.5
K123	459	377.22	30.9
V125	596	887.82	10.2

Table 8.7 k_{ex} and $\Delta\omega$ values obtained for the residues of cL-BABP/S-S complexed with GCDA at a 1:0.1 P:L ratio, showing a RD profile. The obtained parameter χ^2, indicative of the goodness of the fit, is reported in the third column.

cL-BABP/S-S GCDA 1/1	$k_{ex}(s^{-1})$	$\Delta\omega$(Hz)	χ^2
E12	694.8	507.33	19.9
E15	35.4	60.25	7.2
L18	599.3	66.33	11.4
A20	8.3	383.89	5.9
L21	4.6	838.78	5
L23	6.2	483.52	4.7
L27	3.6	585.46	2.4
I28	1889.3	843.46	9.6
M30	104.5	356.24	4
A31	9.1	788.85	4.5
T72	187	492.92	26.1
D74	850.2	175.14	53
K79	557.6	648.97	2.4
L84	392.8	43.43	4.6
L89	808.6	105.83	20.5
S93	6.6	838.96	9.3
V125	1213.5	172.14	88

Table 8.8 k_{ex} and $\Delta\omega$ values obtained for the residues of cL-BABP/S-S complexed with GCDA at a 1:1 P:L ratio, showing a RD profile. The obtained parameter χ^2, indicative of the goodness of the fit, is reported in the third column.

cL-BABP/S-S GCDA 1/ 2.5	$k_{ex}(s^{-1})$	$\Delta\omega$(Hz)	χ^2
D46	4.9	946.13	7.9
G114	10.5	360.5	5

Table 8.9 k_{ex} and $\Delta\omega$ values obtained for the residues of cL-BABP/S-S complexed with GCDA at a 1:2.5 P:L ratio, showing a RD profile. The obtained parameter χ^2, indicative of the goodness of the fit, is reported in the third column.

8.3 Mass Spectrometry results

To confirm NMR data on cL-BABP/S-S binding stoichiometry and to investigate the occupancy of the two binding sites of the protein, mass spectra of cL-BABP/S-S at different ratios of ^{15}N GCDA were acquired.

The obtained data confirmed that cL-BABPs, even in the presence of the disulphide bridge, maintains the ability to bind two GCDA molecules.

Moreover, mass spectrometry analysis of cL-BABP/S-S complexed with GCDA indicated that both the singly ligated (PL_1) and the doubly ligated (PL_2) forms are present at different concentrations of the ligand (See Figure 8.8), in agreement with the general binding mechanism obtained from NMR data.

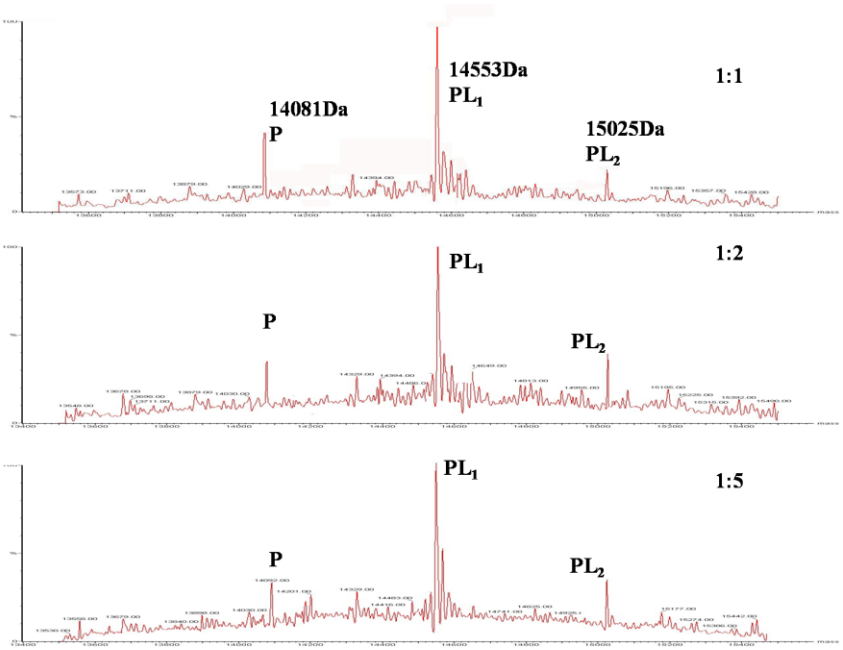


Figure 8.8 Mass spectra of cL-BABP/S-S complexed with ^{15}N GCDA at different protein:ligand (1:1, 1:2, 1:5) ratios. P indicates the signal corresponding to the apo protein, PL_1 to the singly ligated form, and PL_2 to the doubly ligated form.

Table 8.10 reports the intensity percentages of the peaks for 1:1, 1:2 and 1:5 protein:ligands ratios.

	Intensity percentage(%)		
	P	**PL_1**	**PL_2**
15N GCDA 1:1	24	64	12
15N GCDA 1:2	22	63	15
15N GCDA 1:5	13	52	35

Table 8.10 Percentages of the peak intensities of P, PL_1 and PL_2 at 1:1, 1:2 and 1:5, protein:ligand ratios, where P represent the apo cL-BABP/S-S; PL_1 the singly ligated form; PL_2 the doubly ligated form.

Table 8.11 reports the normalized values of the intensities scaled on P values.

	Normalised Intensity(%)		
	P	**PL_1**	**PL_2**
15N GCDA 1:1	1	2.7	0.5
15N GCDA 1:2	1	2.9	0.7
15N GCDA 1:5	1	2.7	2.3

Table 8.11 Normalized values of P, PL_1 and PL_2 relative intensities, scaled on P, for 1:1, 1:2 and 1:5, protein:ligand ratios, where P represent the apo cL-BABP/S-S; PL_1 the singly ligated form; PL_2 the doubly ligated form.

9. Conclusions

The structural data, already known for chicken Liver Bile Acid Binding Protein (cL-BABP) devoid of disulphide bridge, and those obtained in the present work for the protein with a disulphide bridge (cL-BABP/S-S), represented the body of starting information necessary for further analysis of the binding kinetic characteristic of these proteins. The dynamics and the mechanism of protein ligand interactions are strongly related to BABPs biological functions. A number of factors have been shown to modulate ligand binding, e.g. chemistry of ligand and nature of protein residues but, here, the attention was mainly focused on another source of ligand binding variability: the presence of disulphide bridge (S-S). In intracellular proteins, disulfide bonds are generally transiently formed, owing to the reductive nature of the cellular environment. It's plausible that equilibrium between oxidized and reduced forms of cL-BABPs may respond to different needs for ligand uptake.

NMR titration experiments, recorded to observe alternatively the enriched bile acids and the proteins, have shown that the presence of a disulphide bridge, while maintaining the same binding stoichiometry, induces changes in binding ability, site selectivity and dynamic properties of cL-BABPs. Both NMR analysis and mass spectrometry data confirmed the ability of cL-BABP/S-S and cL-BABP to bind two GCDA or GCA molecules.

The most relevant feature emerging from NMR data analysis is the ability of the disulphide bridge to modulate recognition at both sites. No site selectivity for the two bile salts, GCDA and GCA, was previously observed for cL-BABP devoid of disulphide bridge. It is now clear that when the disulphide bridge is present, GCDA preferentially binds to the superficial site and GCA to the internal one. Site selectivity was however observed only when GCDA and GCA are present simultaneously, suggesting that the observed selectivity does not derive from steric exclusion of one bile salt from a specific site.

Chemical Shift Perturbation (CSP) data and cL-BABP/S-S TALOS analysis indicated that no significant change in the tridimensional structure of the apo protein occurred upon disulphide bridge addition. NMR data indicated only minor conformational changes gathered around the inner binding site.

Significant changes in backbone motions were, at variance, observed for several residues distributed over the entire protein. Indeed the presence of the S-S bridge favoured the propagation of slow motions from the C-terminal region of the molecule to the N-terminal β-sheet in the apo protein, and enhanced backbone motions in cl-BABP/S-S-CDA complex, differently to what observed for cL-BABP, where the binding of this ligand was accompanied by a substantial quench of motions.

In conclusion site-selectivity appeared to be governed by protein mobility, rather than by steric factors related to the hydroxylation pattern of the ligand. These results underlined the tight connection between ligand binding phenomena and protein mobility in this protein family and set the basis for NMR kinetic studies based on lineshape analysis and relaxation dispersion measurements.

NMR lineshapes were analysed as a function of ligand concentration and simulated with the NMRKIN software specifically implemented to treat multisite binding mechanisms.

The ligand binding mechanism, investigated through the combination of lineshape analysis and relaxation dispersion experiments, allowed the distinction between two fundamentally different binding mechanisms namely, "induced fit" versus "conformational selection". The binding mechanism resulted extremely complex to be deconvoluted, but the reported data unambiguously support the "conformational selection" model for cL-BABPs. The apo protein samples both the competent and non-competent binding conformations, involving opening at the portal area. Despite the complexity exhibited by ternary complex, the mechanism involved in the binding of the two ligands could be tackled providing an estimate of the kinetics involved. The presented lineshape analysis allowed us to rationalise the behaviours observed for different residues in the two proteins (cL-BABP/S-S and cL-BABP) contributing to the definition of the following simplified multi-step binding mechanism:

$$P \rightleftarrows P^* + L \rightleftarrows PL_n + L \rightleftarrows PL_1L_2 \rightleftarrows PL_1L_2^* \quad \text{with } n=1,2$$

where P represents the apo protein; P* represents a possible low-populated conformation visited by the apo protein; L represents the ligand; PL_n represents a generic intermediate which could derive either from differences in protein conformations or in the type of bound species; PL_1L_2 represents the protein at saturation and PL_1L_2* is a low-populated conformation of the holo protein.

The obtained results showed that the presence/absence of a disulphide bridge in L-BABPs is not affecting the general binding mechanism, rather it seems to have an

influence on the kinetics of the first exchange process. In general cL-BABP exhibits lower $k_{off,1}$ rates compared to cL-BABP/S-S suggesting a higher affinity of cL-BABP for the ligand. Slow (in the range of 20-60 s^{-1}) or fast to intermediate(~ 100-2000 s^{-1}) $k_{off,1}$ rates were deduced from lineshape simulations of cL-BABP residues. Fast to intermediate (~ 400-600 s-1) or fast (≥ 2000 s-1) $k_{off,1}$ rates were deduced for cL-BABP/S-S, which in turn translates, into a less cooperative but site-selective system towards GCDA. cL-BABP does not display any site-selectivity, but is an extremely cooperative system. Thus the different behaviour of the two proteins may be related to the bile-acid binding capability, which seems to be directly correlated with cooperativity.

These findings represent the first step towards the global description of the mechanism of bile acid binding and release within the cell, necessary to understand the role of L-BABPs and their involvement in cholesterol homeostasis. The gain of this research could be relevant, among others, in the field of health. Indeed the understanding of mechanisms regulating the interaction of bile acids with their cytoplasmic transporters is necessary for the design of drugs which may have utility in treating metabolic dysfunctions.

Tables

Table 1 Backbone (**^{1}**H, ^{15}N) and side chain (Cα and Cβ) assignments of BABP/S-S apo form, pH 7.2, 298K

Residue	HN	N	Cα	Cβ
A1	8,71	125,27	51,50	19,09
F2	8,40	113,00	58,52	38,22
S3	8,00	111,19	57,73	61,59
G4	9,06	111,36	41,85	
T5	8,36	117,70	60,37	66,90
W6	9,50	127,67	52,84	28,33
Q7	9,40	123,50	51,35	28,60
V8	8,31	131,32	60,96	29,13
Y9	9,52	124,01	54,13	38,03
A10	7,37	125,76	49,35	19,86
Q11	8,92	120,41	52,20	28,70
E12	9,27	124,85	53,81	28,60
N13	9,23	118,45	51,58	35,38
Y14	7,89	118,01	56,61	36,40
E15	8,99	118,00	58,65	26,52
E16	8,68	118,02	57,50	26,16
F17	8,23	123,39	59,92	37,83
L18	8,53	117,89	55,54	39,30
K19	7,96	117,28	57,10	29,68
A20	8,00	125,71	52,43	15,30
L21	7,30	116,87	52,70	39,79
A22	8,00	118,26	50,12	13,49
L23	7,13	120,01	51,10	39,23
P24			60,03	30,07
E25	8,83	122,30	57,70	27,27
D26	8,92	116,88	54,40	36,98
L27	7,28	120,49	54,50	39,02
I28	7,89	121,49	63,85	35,02
K29	7,85	117,70	57,10	29,80
M30	7,25	115,29	54,88	30,76
A31	8,05	121,21	51,34	16,95
R32	7,97	115,95	56,80	27,23
D33	8,05	117,27	52,10	39,28
I34	7,63	121,13	58,44	35,70
K35	8,76	129,10	50,52	30,05
P36			60,17	30,59
I37	8,10	118,48	57,89	37,97
V38	9,60	128,89	58,50	32,13
E39	9,64	128,24	52,25	29,82
I40	9,57	125,50	58,62	37,57

Q41	9,24	127,69	52,42	29,15
Q42	8,31	125,54	51,71	27,87
K43	8,31	130,17	52,85	30,79
G44	8,94	117,30	45,01	
D45	8,85	127,51	52,08	38,92
D46	8,01	119,87	51,74	39,33
F47	9,28	121,40	54,23	41,17
V48	8,57	121,66	60,27	31,91
V49	9,31	128,04	59,02	31,94
T50	9,74	127,77	59,15	68,33
S51	9,37	122,73	54,60	62,12
K52	9,64	126,07	53,96	33,33
T53	8,52	115,98	56,70	67,11
P54			63,19	29,58
R55	7,97	111,49	53,76	29,39
Q56	7,60	117,77	52,51	29,87
T57	8,70	123,26	59,89	68,36
V58	8,90	127,02	59,29	32,65
T59	8,96	125,52	59,31	68,20
N60	9,24	125,30	50,14	40,69
S61	8,81	116,89	54,43	62,69
F62	8,34	119,00	53,97	38,08
T63	9,24	119,01	58,73	69,91
L64	9,07	125,00	54,32	39,26
G65	8,74	108,50	43,21	
K66	7,90	119,93	51,86	32,46
E67	8,69	125,10	55,82	28,25
A68	9,18	130,54	48,62	20,37
D69	8,26	119,57	51,12	39,55
I70	8,99	125,31	55,93	37,08
T71	9,12	122,47	59,51	67,12
T72				
M73			56,27	30,47
D74	8,19	114,51	50,83	38,56
G75	8,06	107,75	43,15	
K76	7,42	121,30	52,95	30,23
K77				
L78			51,41	42,14
K79	8,12	120,21	52,24	32,70
C80	8,69	116,83	52,26	43,009
T81	8,78	119,14	60,33	66,49
V82	8,68	131,26	59,26	29,38
H83	8,61	123,73	52,71	31,12
L84	9,09	123,97	51,37	

A85	9,02	128,64	49,02	
N86			51,67	
G87	8,62	103,30	43,23	
K88	7,87	120,32	52,70	31,80
L89			52,64	42,36
V90	9,21	127,60	52,64	31,89
C91	8,50	125,50	51,68	41,21
K92	8,86	123,90	53,09	33,67
S93	9,19	121,27	54,20	63,80
E94	8,99	120,70		
K95				
F96	8,88	116,90	54,30	41,60
S97	8,12	119,41	58,90	67,60
H98	8,00	123,90	60,60	29,90
E99	8,25	121,80	58,40	36,40
Q100	8,37	127,60	51,50	30,15
E101	8,66	123,73	52,44	31,40
V102	8,70	126,50	59,17	31,34
K103	8,95	129,26	52,66	31,14
G104	9,11	116,74	45,15	
N105	8,98	125,50	50,68	36,45
E106	8,09	120,34	53,18	30,94
M107	8,80	123,77	51,65	32,73
V108	8,95	126,50	59,10	31,87
E109	9,26	128,70	51,30	31,80
T110	8,75	120,13	59,40	66,90
I111	9,44	132,25	58,13	37,47
T112	9,29	122,73	58,46	68,43
F113	9,04	125,26	55,32	40,49
G114	8,46	115,50	44,44	
G115	8,50	105,57	42,50	
V116	8,02	123,36	59,87	31,02
T117	8,72	123,53	59,82	67,86
L118	9,60	132,50	50,92	41,88
I119	7,85	125,25	56,96	36,85
R120	9,24	126,76	52,33	32,00
R121	8,93	126,51	52,65	29,64
S122	9,33	117,75	55,52	65,02
K123	8,45	118,0	51,45	33,82
R124	8,30	125,73	54,65	27,94
V125	8,31	129,53	61,37	30,62

Table 2 Backbone (**^{1}H, ^{15}N**) and side chain (Cα and Cβ) assignments of BABP/S-S complexed with CDA (P:L ratio 1:3), pH 7.2, 298K

Residue	HN	N	Cα	Cβ
A1			51,67	16,51
F2	8,55	113,37	58,20	38,04
S3	7,93	111,63	58,67	61,32
G4	9,07	111,07	41,57	
T5	8,21	118,00	59,96	66,40
W6	9,61	127,57	52,65	29,51
Q7			51,84	29,53
V8	8,24	131,82	61,04	
Y9	9,55	123,85	54,16	38,18
A10	7,33	125,98	49,22	19,28
Q11	8,72	119,20	51,86	28,32
E12	9,09	123,85	53,14	28,89
N13	9,11	118,00	51,53	34,26
Y14	7,80	118,53	57,90	36,68
E15	8,95	118,53	58,12	26,25
E16	9,03	118,48	58,03	26,07
F17	8,36	123,85	59,88	37,69
L18	8,27	116,94	55,55	37,92
K19	7,79	116,64	56,30	29,54
A20	7,95	126,50	52,35	14,96
L21	7,37	116,94	52,53	40,12
A22	7,83	118,53	49,81	13,15
L23	6,87	118,48	50,96	37,94
P24			59,83	29,66
E25	8,82	123,32	57,51	26,78
D26	8,89	116,41	54,27	36,98
L27	7,19	122,22	54,42	39,25
I28	7,87	120,11	63,89	34,67
K29	7,80	116,94	57,05	29,82
M30	7,44	116,94	55,07	30,13
A31	8,55	120,66	50,583	16,58
R32	7,77	121,18	57,85	27,76
D33	8,24	116,40	57,89	39,11
I34	7,31	120,34	58,00	35,27
K35	8,92	131,29	50,19	29,97
P36				
I37			57,48	37,90
V38	9,45	129,70	58,03	37,72
E39	9,56	128,63	52,01	29,44
I40	9,63	125,98	58,31	37,38
Q41	9,25	128,10	52,02	28,67

Q42	8,23	126,50	51,10	27,56
K43	8,80	130,29	52,69	30,46
G44	8,92	118,00	44,82	
D45	8,84	128,09	51,84	38,53
D46	8,01	120,66	51,70	38,86
F47	9,27	122,25	53,86	41,00
V48	8,52	121,72	59,92	31,90
V49	9,31	127,57	58,42	31,92
T50	9,80	129,16	58,88	67,89
S51	9,15	122,25	54,09	61,68
K52	9,34	126,50	53,28	33,48
T53	8,97	116,67	56,26	66,65
P54			63,10	30,11
R55	8,24	112,70	54,22	30,11
Q56	7,68	116,40	52,70	30,17
T57	8,49	118,00	58,99	68,68
V58	8,90	125,97	58,75	32,56
T59	8,96	127,03	59,35	67,94
N60	9,13	124,90	50,35	41,45
S61	8,85	116,40	54,16	62,83
F62	8,25	118,00	54,01	37,53
T63	9,65	120,13	58,29	69,10
L64	9,21	127,03	53,62	38,62
G65	9,01	107,38	43,09	
K66	7,92	121,19	51,66	31,87
E67	8,92	130,22	56,44	27,73
A68	8,67	131,29	48,27	19,62
D69	8,32	120,13	52,01	39,21
I70	8,95	125,80	54,82	36,66
T71	9,41	123,85	58,44	67,54
T72	8,55	115,88	57,83	68,23
M73	10,32	120,13	56,21	31,81
D74	8,68	112,70	50,26	35,93
G75	7,57	107,90	43,57	
K76	8,12	120,67	53,18	30,93
K77	8,04	119,60	52,01	30,59
L78	9,03	123,85	50,76	42,84
K79	8,21	119,60	52,07	33,75
C80	8,39	117,00	54,04	44,60
T81	8,74	124,00	59,61	65,34
V82	9,51	132,35	58,92	28,54
H83	8,65	123,85	52,34	30,92
L84	8,77	122,25	50,92	41,48
A85	9,24	129,70	48,70	18,07

N86			51,41	34,57
G87	8,58	103,13	42,98	
K88	7,83	120,13	52,83	31,75
L89	8,97	124,90	51,79	41,07
V90	8,96	121,72	57,82	33,28
C91	8,86	125,40	58,4	39,93
K92	9,23	128,60	52,50	32,99
S93	8,49	119,60	52,76	63,36
E94	9,15	128,78	56,21	26,69
K95	8,15	114,81	53,00	32,26
F96	6,93	113,22	52,98	40,34
S97	8,80	112,59	54,80	60,87
H98	9,16	129,06	54,55	32,99
E99	7,99	125,10	52,09	31,62
Q100				
E101	8,87	125,052	52,00	32,15
V102	8,73	124,91	58,83	31,54
K103	8,92	129,16	52,38	31,04
G104	9,40	115,53	44,97	
N105			50,33	35,93
E106	8,06	120,13	52,79	30,59
M107	8,87	123,95	51,09	33,77
V108	9,02	126,50	58,74	32,73
E109	9,11	129,16	51,66	31,40
T110	8,84	120,67	59,14	66,74
I111	9,35	130,22	57,50	37,40
T112	9,47	122,79	58,38	68,26
F113	9,12	127,05	55,88	40,95
G114	8,36	116,94	44,55	
G115	8,53	105,25	42,26	
V116	8,14	124,38	59,55	30,98
T117	8,74	123,46	59,41	66,85
L118	9,57	132,35	50,68	41,64
I119	7,73	124,91	56,88	36,84
R120	9,42	127,57	52,41	31,85
R121	9,09	127,03	52,48	29,13
S122	9,28	118,21	55,49	63,10
K123	8,66	120,67	51,06	33,54
R124	8,17	124,91	54,24	27,81
V125	8,17	128,96	60,96	30,50

Table 3 Backbone (**^{1}**H, ^{15}N) assignments of BABP/S-S complexed with GCDA (P:L ratio 1:3), pH 7.2, 298K

Residue	HN	N
A1		
F2		
S3	7,95	111,40
G4	9,05	111,45
T5	8,18	117,77
W6	9,61	127,50
Q7	9,20	120,90
V8	8,26	131,44
Y9	9,52	123,13
A10	7,32	125,87
Q11	8,70	118,34
E12	9,09	123,37
N13	9,13	117,64
Y14	7,99	118,48
E15	8,94	118,51
E16	9,09	118,80
F17	8,41	124,11
L18	8,30	117,01
K19	7,78	117,05
A20	7,93	126,14
L21	7,43	116,47
A22	7,91	118,79
L23	6,94	118,64
P24		
E25	8,83	123,24
D26	8,93	116,60
L27	7,25	121,69
I28		
K29	7,69	116,53
M30	7,38	117,04
A31	8,49	120,35
R32	7,83	120,89
D33	8,09	116,08
I34	7,27	120,34
K35		
P36		
I37		
V38	9,44	129,01
E39	9,51	128,32
I40	9,65	126,13
Q41	9,24	128,17

Q42	8,24	125,97
K43	8,82	130,39
G44	8,93	117,57
D45	8,85	127,56
D46	8,03	120,38
F47	9,27	122,38
V48	8,52	121,76
V49	9,34	127,08
T50	9,71	128,49
S51	9,16	122,46
K52	9,40	125,93
T53	8,87	116,19
P54		
R55	7,98	111,33
Q56	7,57	118,62
T57	8,60	120,83
V58	8,86	125,63
T59	8,95	126,51
N60	9,17	124,67
S61	8,83	116,20
F62	8,26	117,56
T63	8,64	119,89
L64	9,23	126,60
G65	9,07	107,52
K66	7,92	121,13
E67	8,99	130,34
A68	8,72	130,92
D69	8,32	120,45
I70	8,90	125,93
T71	9,39	123,99
T72	8,63	116,02
M73	10,43	120,16
D74	8,70	112,11
G75	7,59	107,88
K76	8,13	120,29
K77	7,91	120,74
L78	9,01	124,05
K79	8,23	119,32
C80	8,27	116,37
T81	8,78	122,41
V82	9,48	131,95
H83	8,67	123,45
L84	8,82	122,37
A85	9,25	129,25

N86		
G87	8,60	103,33
K88	7,86	120,52
L89		
V90	8,93	121,63
C91	8,88	124,74
K92		
S93	8,55	119,07
E94		
K95	8,31	114,10
F96	7,03	113,53
S97	8,80	112,44
H98	9,10	127,85
E99	7,89	125,13
Q100	9,08	125,06
E101	8,91	125,41
V102	8,77	125,05
K103	8,92	129,31
G104	9,11	117,56
N105		
E106	8,09	120,05
M107	8,90	123,95
V108	9,05	126,91
E109	9,29	128,18
T110	8,72	120,75
I111	9,41	130,23
T112	9,51	122,77
F113	9,03	125,75
G114	8,43	116,10
G115	8,58	105,36
V116	8,09	123,87
T117	8,74	123,43
L118	9,65	132,55
I119	7,79	125,11
R120	9,36	126,12
R121	9,14	127,02
S122	9,31	118,32
K123	8,63	120,41
R124	8,10	124,49
V125	8,19	128,16

Table 4 Backbone (**^{1}**H, ^{15}N, ^{13}C) and side chain (Cα and Cβ) assignments of BABP/S-S complexed with GCDA-GCA (P:L ratio 1:2:2), pH 7.2, 298K

Residue	HN	N	Cα	Cβ	C
A1					
F2			61,17	40,69	175,90
S3	7,94	111,31	59,97	64,14	174,70
G4	9,06	111,53	44,30		171,80
T5	8,21	117,66	62,72	69,31	173,190
W6	9,53	127,77	55,49	32,29	
Q7			54,55	32,27	175,90
V8	8,28	131,58	63,60	32,20	
Y9	9,54	123,32	56,80	40,88	174,60
A10	7,34	125,87	51,98	22,16	175,20
Q11	8,69	118,37	54,46	32,87	174,70
E12	9,12	123,79	55,55	31,76	175,70
N13	9,13	117,55	54,10	36,80	175,60
Y14	7,89	118,28	60,70	39,45	176,50
E15	8,95	118,61	60,89	28,70	178,50
E16	9,10	118,79	60,30	28,60	178,60
F17	8,41	124,24	62,60	40,38	176,40
L18	8,20	116,41	58,34	40,75	178,80
K19	7,78	116,59	59,07	32,15	179,80
A20	7,97	125,98	55,04	17,80	178,60
L21	7,37	116,86	55,40	42,70	176,10
A22	7,92	118,82	52,40	15,96	177,00
L23	6,95	118,47	53,60	41,11	174,50
P24			62,50	32,60	177,50
E25	8,81	122,50	60,10	29,69	177,90
D26	8,92	116,62	57,08	39,62	177,90
L27	7,22	121,99	57,16	42,23	178,07
I28	7,88	120,35	66,59	37,29	177,40
K29	7,84	117,01	59,78	32,39	178,90
M30	7,44	117,17	57,70	32,97	177,50
A31	8,41	120,25	53,19	19,29	
R32	7,75	121,91	60,70	30,60	176,50
D33	8,35	116,23	54,22	41,76	175,90
I34	7,22	120,06	60,80	37,86	176,10
K35	8,99	130,99	52,97	32,75	173,70
P36					
I37			60,26	40,17	175,50
V38	9,54	129,02	61,13	34,85	175,10
E39	9,59	128,35	54,54	32,07	176
I40	9,63	125,98	61,33	40,16	175,80
Q41	9,25	128,16	54,94	31,36	173,90

Q42	8,24	125,94	54,15	30,25	174,70
K43	8,81	130,36	55,26	33,27	176,03
G44	8,94	117,52	47,48		174,30
D45	8,86	127,51	54,57	41,27	174,70
D46	8,08	119,68	54,29	41,57	175,07
F47	9,28	122,53	56,42	43,80	176,40
V48	8,52	121,69	62,77	34,77	175,70
V49	9,31	127,22	61,18	34,66	175,70
T50	9,73	128,52	61,53	70,76	173,81
S51	9,17	122,59	56,99	64,55	173,80
K52	9,46	125,38	56,12	36,13	175,90
T53	8,73	116,01	59,14	68,78	173,40
P54			66,08	32,08	177,20
R55	8,00	111,52	55,60	32,52	175,80
Q56	7,52	118,61	55,61	33,09	174,20
T57	8,59	121,60	61,50	71,50	174,00
V58	8,82	125,49	61,38	35,56	174,40
T59	8,99	126,57	61,98	70,77	173,40
N60	9,18	124,70	53,00	44,27	173,30
S61	8,84	116,25	56,91	65,59	173,30
F62	8,23	117,63	56,59	40,03	173,10
T63	9,64	119,92	61,10	72,06	175,60
L64	9,26	126,68	56,10	41,33	178,00
G65	9,08	107,52	45,78		173,80
K66	7,91	121,07	54,44	34,62	175,60
E67	8,92	128,89	59,22	30,40	175,03
A68	8,64	130,94	51,00	22,10	176,00
D69	8,33	120,50	54,67	41,93	174,40
I70	8,83	125,44	57,69	39,67	175,30
T71	9,44	123,95	61,61	70,37	174,60
T72	8,69	116,15	60,55	70,71	177,30
M73	10,40	119,90	58,90	34,45	176,20
D74	8,65	112,23	52,63	38,42	175,60
G75	7,59	108,04	46,40		175,00
K76	8,25	120,45	54,24	33,50	175,40
K77	8,12	119,47	54,36	32,84	176,10
L78	9,01	123,77	53,17	45,78	175,60
K79	8,23	119,28	55,00	36,85	176,60
C80	8,29	116,27	56,50	45,70	173,50
T81	8,84	122,44	62,46	67,85	173,10
V82	9,48	131,83	61,61	31,20	175,10
H83	8,68	123,47	54,99	33,40	173,50
L84	8,84	122,03	53,79	44,31	176,50
A85	9,22	128,80	55,48	20,91	177,00
N86			54,14	37,55	175,40

G87	8,62	103,43	45,67		173,80
K88	7,85	120,74	54,14	34,61	175,50
L89	8,77	124,60	54,50	43,70	176,70
V90	9,03	122,66	60,35	35,63	175,20
C91	8,82	123,42	55,23	42,97	172,70
K92	9,44	128,97	55,36	35,96	174,27
S93	8,56	118,78	55,43	66,07	174,20
E94	8,60	119,46	59,44	29,36	177,60
K95	8,29	113,91	56,18	35,22	175,98
F96	6,98	113,34	55,71	43,56	173,90
S97	8,78	112,21	57,87	65,99	172,56
H98	9,16	126,31	56,26	36,27	174,50
E99	8,31	126,82	55,00	34,00	175,40
Q100					
E101			55,29	34,80	
V102	8,77	124,83	61,38	34,10	175,20
K103	8,94	129,42	55,34	33,76	176,50
G104	9,34	119,14	47,68		174,50
N105	9,01	125,35	53,13	39,06	
E106	8,10	120,08	55,55	33,16	
M107	8,91	124,08	53,70	36,80	
V108	9,017	126,82	61,34	34,90	
E109	9,10	127,91	54,44	34,27	174,41
T110	8,85	120,76	61,98	69,52	174,90
I111	9,47	131,56	60,31	40,78	174,60
T112	9,46	122,80	61,37	71,25	174,37
F113	9,00	124,67	58,80	43,48	175,70
G114	8,41	116,25	46,84		175,30
G115	8,57	105,64	44,89		174,20
V116	8,15	124,05	62,25	33,82	174,64
T117	8,74	123,47	62,17	70,45	173,80
L118	9,69	132,76	53,57	44,73	175,20
I119	7,79	125,02	59,34	39,55	175,90
R120	9,37	126,27	54,99	34,13	175,30
R121	9,09	126,82	55,25	31,83	175,02
S122	9,27	118,30	58,10	66,85	173,01
K123	8,62	120,15	53,93	36,33	175,60
R124	8,18	124,87	56,90	30,38	176,50
V125	8,22	128,52	63,57	33,18	

Appendices

```
% ========================================================================
%              FITFTEX - Main program to fit lineshapes.
% ========================================================================

% 2009 version with parameter handling through a structure.
% A parameter file similar to testpar.m is required.
%
% Needs: A parameter file (sample file: testpar.m)
%        a routine which creates the parameter structure (mklspar.m)
%        data matrices
%        uses redat.m to read data from disk (if necessary)
%        a mechanism file (e.g. m21)
%

% Copyright (c)
% Ulrich Guenther 20-3-97, Feb 2008
%

global SPC;             % spectra parameters
global DISP;            % display parameters
global LSPAR;           % Lineshape parameters
global HANDLES;         % graphics handles
global P;

%plot_LW1 = 1.5;
%plot_LW2 = 3.0;
plot_LW1 = 0.5;
plot_LW2 = 2.0;
MarkSiz  = 9;
MarkLW   = 2.8;

DEBUG=0;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%filename = 'testpar.m';
%filename = 'testpar_m21.m';
filename = 'testpar_m2121.m';
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear FN
LFN=length(filename);
FN(1:LFN-2)=filename(1:LFN-2);

LSPAR=mklspar2(FN);         % no path for feval in mklspar

% ************************************************
% ****************** READ DATA *******************
% ************************************************
if LSPAR(1).USE_DATA
  disp('Reading dim 1 data.')
  % DIM 1 setup

  if isempty(LSPAR(1).w) | LSPAR(1).w==0
    % Call graphical routine to set maxima of peaks
    MX1 = redat2(1,1);              % dim 1, set w
  else
    MX1 = redat2(1,0);          % dim1, don't set w
  end
  [MX1siz no_spcs_1] = size(MX1);
  LSPAR(1).MXsiz     = MX1siz;
```

```
  LSPAR(1).NO_SPCS   = no_spcs_1;
  no_spcs            = no_spcs_1;

  % ************************************************************
  % ********************** READ DIM 2 DATA **********************
  % ************************************************************
  if LSPAR(1).DIM==2 & ~isempty(LSPAR(2).filename)
    disp('Reading dim 2 data.')
    if isempty(LSPAR(1).w) | LSPAR(1).w==0
      MX2 = redat2(2,1);
    else
      MX2 = redat2(2,0);
    end
    [MX2siz no_spcs_2] = size(MX2);
    LSPAR(2).MXsiz     =  MX2siz;
    if no_spcs_2 ~= no_spcs_1
      error('Number of spectra in dim 1 and 2 differs.')
    end
  end
  PLOT_DIM_2_DATA=1;
else    % no DIM 2!
  MX2siz  = LSPAR(2).MXsiz;
  MX1siz  = LSPAR(1).MXsiz;
  no_spcs = LSPAR(1).NO_SPCS;
  PLOT_DIM_2_DATA = 0;
end

% Create Cell array MX
clear MX
if LSPAR(1).USE_DATA & LSPAR(1).DIM==1
  temp=MX1;
  MX{1}=temp;
elseif LSPAR(1).USE_DATA & LSPAR(1).DIM==2
  temp=MX1;
  MX{1}=temp;
  MX{2}=MX2;
  %  MX(:,:,1) = MX1;
  %  MX(:,:,2) = MX2;
elseif LSPAR(1).USE_DATA
  error('DIM must be 1 or 2!')
end

% call mechanism file, 2nd dim is calles autorecursively in mechanism file
if exist(LSPAR(1).MECH)==2              % MECH must exist as .m-file
  if LSPAR(1).DIM==2 & exist(LSPAR(2).MECH)==2
    disp('++++++++++++++++++++++++++++++++++++++++');
    disp('     Calculating 2D lineshapes');
    disp('++++++++++++++++++++++++++++++++++++++++')
    [calcM,calcXY] = feval(LSPAR(1).MECH,[],'CALC',1);

  elseif LSPAR(1).DIM==1
    disp('++++++++++++++++++++++++++++++++++++++++');
    disp(['     Calculating dim ' num2str(LSPAR(1).DIM) ]);
    disp('++++++++++++++++++++++++++++++++++++++++')
    calcM = feval(LSPAR(1).MECH,[],'CALC',1);
    % Convert to cell array for compatibility with 2DIM
    temp = calcM;
    clear calcM
    calcM{1} = temp;
    clear temp
```

```
  else
    error('Probably some mechanism missing.')
  end
else
  tmp_str=['DIM 1: Mechanism ' LSPAR(1).MECH ' does not exist.'];
  error(tmp_str)
  return
end

% ****************** Set display pars ******************
if ~isfield(LSPAR(1),'colstep')
  colstep = 1/no_spcs;
  for n=1:no_spcs
    % CyanRed
    % ColPal(n,:) = [n*colstep, 1-n*colstep, 1-n*colstep];
    % BlueRed
    ColPal(n,:) = [n*colstep, 0.0, 1-n*colstep];
  end
  LSPAR(1).ColPal = ColPal;
  LSPAR(2).ColPal = ColPal;
end

if LSPAR(1).USE_DATA==0
  MX = [];
end
if LSPAR(1).DIM==1
  calcXY = [];
end

plot_lineshapes(MX,calcM, calcXY)

if LSPAR(2).realdim==1
  LSPAR = LSPAR([2,1]);
end % if ~(LSPAR(1).DIM2MODE==2 & DIM==2)


% ***********************************************************
%                         Data fitting
% ***********************************************************

FIT=LSPAR(1).FIT;

if FIT
  clear OPTIONS
  OPTIONS=optimset;
  % OPTIONS=optimset(OPTIONS,'HessUpdate','dfp');
  %OPTIONS=optimset(OPTIONS,'HessUpdate','steepdesc');
  %OPTIONS=optimset(OPTIONS,'HessUpdate','dfp');
  %OPTIONS=optimset(OPTIONS,'HessUpdate','bfgs');
  %OPTIONS.MaxIter = 5000;
  %OPTIONS.maxFunEvals = 10000;
  OPTIONS=optimset(OPTIONS,'maxFunEvals',5000);
  OPTIONS=optimset(OPTIONS,'tolX',1e-12);
  OPTIONS=optimset(OPTIONS,'tolFun',1e-6);

  OPTIONS=optimset(OPTIONS,'Algorithm','trust-region-reflective');
  %OPTIONS=optimset(OPTIONS,'Algorithm','levenberg-marquardt');

  %OPTIONS=optimset(OPTIONS,'LevenbergMarquardt','on');
```

```
  %OPTIONS=optimset(OPTIONS,'GaussNewton','on');
  %OPTIONS=optimset(OPTIONS,'LevenbergMarquardt','off');
  %OPTIONS=optimset(OPTIONS,'LargeScale','off');

  PAR_0 = feval(LSPAR(1).MECH,[],'PAR0');
  par0  = PAR_0{1};
  LB    = PAR_0{2};
  UB    = PAR_0[85];

  %PAR_0 = m21_2D([],'PAR0');
  %PAR_0 = m2121_2D([],'PAR0');

  %[par_final,fval,exitflag,output] =
fminsearch('m21_2D',par0,OPTIONS,'FIT',1,MX);
  %[par_final,fval,exitflag,output] =
fminsearchbnd('m2121_2D',par0,LB,UB,OPTIONS,'FIT',1,MX);
  [par_final,fval,exitflag,output] =
lsqnonlin(LSPAR(1).MECH,par0,LB,UB,OPTIONS,'FIT',1,MX);
end

return
```

```
% ====================================================================
%      m2121 - Mechanism file implemented for cL-BABP/S-S and cL-BABP
%              by U.Günther-2009
% ====================================================================

function [R1, R2] = m2121_2D(par,action,curdim,MX,Plothandle)
%
% m2121 - kinetic mechanism P + L <=> PL_n + L <=> PL_1L_2
%         = both steps are ligand concentration dependent.
%         = i.e. k_on = k_on* [L].
%
% X is:
% X.tau21 = 1;
% X.tau32 = 2;
% X.pa = X.tau21 + (1:no_spcs)';
% X.pb = X.pa    + (1:no_spcs)';

global HANDLES
global LSPAR

% NEW LSPAR type used
%if ndims(LSPAR)==2
%  LSPAR0 = LSPAR;
%  clear LSPAR
%  LSPAR  = LSPAR_conv(LSPAR0);
%end

DISPLAY_PAR=1;
UPDATE_PLOT=1;
VERBOSE=0;

if VERBOSE
  disp(['m2121(' num2str(curdim) ')']);
end

if nargin<1
  error('Usage: [R1, R2, par] = m2121(curdim,MX,par)');
end

% define action: PAR0  - create par0
%                CALC  - calculate calcM and calc FID
%                FIT   - fit data
%

if ~exist('action')
  error('Input argument action may not be empty.')
else
  action=upper(action);
end
if ~exist('curdim')
  curdim=0;
end
if strcmp(action,'FIT') | strcmp(action,'CALC')
  if curdim<1 | curdim>2
    error('curdim must be 1 or 2!')
  end
end
if strcmp(action,'FIT')
  if ~exist('par') | isempty(par)
    error('You must provide par for a successful FIT.')
```

```
  end
  if ~exist('MX') | isempty(MX)
    error('You must provide MX for a successful FIT.')
  end
end
if VERBOSE==1
  disp('#########################')
  disp(['Running m2121 in ' action ' mode.']);
  disp('#########################')
end

% Set global pointers for data fitting
no_spcs = LSPAR(1).NO_SPCS;
X.tau21 = 1;
X.tau32 = 2;
X.pa = X.tau32 + (1:no_spcs)';
X.pb = X.pa    + (1:no_spcs)';
%X.pc = X.pb    + (1:no_spcs)';

% Get parameters for dim2 ************************************
if strcmp(action,'PAR0')
  par0(X.tau21) = LSPAR(1).taus_0(1);
  par0(X.tau32) = LSPAR(1).taus_0(2);
  par0(X.pa)    = LSPAR(1).p_a_0;
  par0(X.pb)    = LSPAR(1).p_b_0;
  % par0(X.pc)    = LSPAR.p_c_0;
  LB(X.tau21)  = 1/200;
  LB(X.tau32)  = 1/20;
  LB(X.pa)     = .0001*ones(1,LSPAR(1).NO_SPCS);
  LB(X.pb)     = .0001*ones(1,LSPAR(1).NO_SPCS);

  UB(X.tau21)  = 1/6000;
  UB(X.tau32)  = 1/5;
  UB(X.pa)     = .9999*ones(1,LSPAR(1).NO_SPCS);
  UB(X.pb)     = .9999*ones(1,LSPAR(1).NO_SPCS);
  PAR_0{1}      = par0;
  PAR_0{2}      = LB;
  PAR_0[85]      = UB;

elseif (strcmp(action,'CALC') | strcmp(action,'FIT')) & curdim==2
  if VERBOSE==1
    disp('#########################')
    disp('  DIM 2 calculation.')
    disp('#########################')
  end

  anz      = LSPAR(1).NO_SPCS;
  no_spcs  = LSPAR(1).NO_SPCS;
  Hzppt    = LSPAR(2).Hzppt;
  MXsiz    = LSPAR(2).MXsiz;
  SWH_MX   = (MXsiz+1) * Hzppt;          % I don't know why +1!
  SWH      = LSPAR(2).SWH;
  TD       = LSPAR(2).TD;
  W        = (LSPAR(2).w-1) * Hzppt;
  wdwf     = LSPAR(2).wdwf;
  E        = ones(1,anz);
  LW       = LSPAR(2).LW;
  zf       = LSPAR(2).zf;
  p_a      = LSPAR(2).p_a_0;
  p_b      = LSPAR(2).p_b_0;
```

```
  %p_c      = LSPAR.p_c_0_2;
  tau21    = LSPAR(2).taus_0(1);
  tau32    = LSPAR(2).taus_0(2);

  % Mechanism specific FIT parameters
  if ~isempty(par)              % Same tau as in dim 1
    par   = par(:)';
    tau21 = par(X.tau21);
    tau32 = par(X.tau32);
    p_a   = par(X.pa);
    p_b   = par(X.pb);
    %p_c   = par(X.pc);
  end

  switch upper(wdwf{1})
   case {'QSINE','SINE'}
    WDWF=wdwf2(ones(MXsiz,1),wdwf{1},wdwf{2},MXsiz);
   case 'GM'
    TF = TD/MXsiz;      % thinning factor
    WDWF=wdwf2(ones(MXsiz,1),wdwf{1},wdwf{2},wdwf[85],SWH/TF);
    disp(['wdwf2(' num2str(wdwf{1}) ',' num2str(wdwf{2}) ',' ...
          num2str(wdwf[85]) ',' num2str(SWH/TF) ')']);
   otherwise
    WDWF = ones(MXsiz,1);
  end

  calcM = ftex2121(LW,tau21,tau32,p_a,p_b,W,MXsiz,SWH_MX, ...
                               WDWF,zf,ones(1,anz));

elseif (strcmp(action,'CALC') | strcmp(action,'FIT')) & curdim==1
  if VERBOSE==1
    disp('DIM 1 calculation.')
  end
  anz        = LSPAR(1).NO_SPCS;
  no_spcs    = LSPAR(1).NO_SPCS;
  Hzppt      = LSPAR(1).Hzppt;
  MXsiz      = LSPAR(1).MXsiz;
  SWH_MX     = (MXsiz+1) * Hzppt;
  SWH        = LSPAR(1).SWH;
  TD         = LSPAR(1).TD;
  W          = (LSPAR(1).w-1) * Hzppt ;
  wdwf       = LSPAR(1).wdwf;
  LW         = LSPAR(1).LW;
  zf         = LSPAR(1).zf;
  E          = ones(1,anz);
  LW         = LSPAR(1).LW;
  zf         = LSPAR(1).zf;
  p_a        = LSPAR(1).p_a_0;
  p_b        = LSPAR(1).p_b_0;
  %p_c        = LSPAR(1).p_c_0;
  tau21      = LSPAR(1).taus_0(1);
  tau32      = LSPAR(1).taus_0(2);
  FITDIM     = LSPAR(1).FITDIM;

  % Mechanism specific FIT parameters
  if ~isempty(par)          % FIT ==> overwrite p_a and tau
    par   = par(:)';
    tau32 = par(X.tau32);
    tau21 = par(X.tau21);
    p_a   = par(X.pa);
```

```
    p_b   = par(X.pb);
  end

  % ****************************************
  %     autorecursive call of 2nd dimension
  % ****************************************

  if LSPAR(1).DIM==2
    if exist(LSPAR(2).MECH)==2
      disp(['Calculating 2nd dim: ' LSPAR(2).MECH])
      % It shouldn't matter now if par is empty
      calcM_2 = feval(LSPAR(2).MECH,par,'CALC',2,[]);
      %max2dim = max(calcM_2)

    else
      tmp_str=['Mechanism ' LSPAR(1).MECH ' does not exist.'];
      disp(tmp_str)
      return
    end

    % Calculate current dimension
    switch upper(wdwf{1})
     case {'QSINE','SINE'}
      WDWF=wdwf2(ones(MXsiz,1),wdwf{1},wdwf{2},MXsiz);
     case 'GM'
      TF = TD/MXsiz;        % thinning factor
      WDWF=wdwf2(ones(MXsiz,1),wdwf{1},wdwf{2},wdwf[85],SWH/TF);
      disp(['wdwf2(' num2str(wdwf{1}) ',' num2str(wdwf{2}) ',' ...
            num2str(wdwf[85]) ',' num2str(SWH/TF) ')']);
     otherwise
      WDWF = ones(MXsiz,1);
    end

    [calcM calcFID]  = ftex2121(LW,tau21,tau32,p_a,p_b,W,MXsiz,SWH_MX, ...
                                WDWF,zf,ones(1,anz));
    [calcMnew,calcXY] = to2D(calcM,calcM_2);

  elseif LSPAR(1).DIM==1
        switch upper(wdwf{1})
     case {'QSINE','SINE'}
      WDWF=wdwf2(ones(MXsiz,1),wdwf{1},wdwf{2},MXsiz);
     case 'GM'
      TF = TD/MXsiz;        % thinning factor
      WDWF=wdwf2(ones(MXsiz,1),wdwf{1},wdwf{2},wdwf[85],SWH/TF);
      disp(['wdwf2(' num2str(wdwf{1}) ',' num2str(wdwf{2}) ',' ...
            num2str(wdwf[85]) ',' num2str(SWH/TF) ')']);
     otherwise
      WDWF = ones(MXsiz,1);
    end
    [calcM calcFID]  = ftex2121(LW,tau21,tau32,p_a,p_b,W,MXsiz,SWH_MX, ...
                                WDWF,zf,ones(1,anz));
  end
end

% ***************** OUTPUT *****************

if strcmp(action,'FIT')
  if LSPAR(1).DIM==2 & curdim==1
    for N=1:2
      D{N} = sum(sum((MX{N}-calcMnew{N}).^2));
```

```
    end
    if FITDIM(1)==1 & FITDIM(2)==1
        R1 = D{1}+D{2};
    elseif FITDIM(1)==0 & FITDIM(2)==1
        R1 = D{2};
    elseif FITDIM(1)==1 & FITDIM(2)==0
        R1 = D{1};
    else
        disp('Don''t know what to do with R1.')
    end
    R2 = [];

  elseif LSPAR(1).DIM==1
    D = sum(sum((MX{1}-calcM).^2));
    R1 = D;
  end

elseif strcmp(action,'CALC')
  if LSPAR(1).DIM==1
    R1=calcM;
    R2 = [];

  elseif LSPAR(1).DIM==2 & curdim==1
    R1=calcMnew;    % cell array!!!
    R2 = calcXY;

  elseif LSPAR(1).DIM==2 & curdim==2
    R1=calcM;
    R2 = [];
  end

elseif strcmp(action,'PAR0')
  R1 = PAR_0;
  clear R2;
end

% ***************************************
%              PLOT output
% ***************************************
if exist('UPDATE_PLOT') & UPDATE_PLOT & strcmp(action,'FIT') & curdim==1
  for DIM=1:LSPAR(1).DIM
    figure(DIM)
    for N=1:no_spcs
      if LSPAR(1).DIM==2
        set(HANDLES.LSSIM{(DIM-1)*no_spcs+N},'ydata',calcMnew[82](:,N),'Color',
LSPAR(1).ColPal(N,:));
        drawnow
      elseif LSPAR(1).DIM==1
        set(HANDLES.LSSIM{(DIM-1)*no_spcs+N},'ydata',calcM(:,N),'Color',
LSPAR(1).ColPal(N,:));
        drawnow
      end
    end
  end
end

%if ndims(LSPAR)==2
%  clear LSPAR
%  LSPAR = LSPAR0;
%end
```

```
return


% ************************************************************************
%  FTEX2121     FTEX2121     FTEX2121     FTEX2121
% ************************************************************************

function [F_w,F_t] = ftex2121(LW,tau21,tau32,p_a,p_b,W,points,SWH,wdwf,zf,scale)
%
% ftex2121.m
% Lineshape calculation for a 2121 model
% calculating the FID for the actual lineshape
% as described by Kubo and Anderson.
%
% Note: expm used for Matrix Exponential, requires square matrix!
% This function can only be called from within m2121!
%

DEBUG=0;
DEBUG1=0;

global LSPAR
global P
if ndims(LSPAR)==2
  LSPAR0 = LSPAR;
  clear LSPAR
  LSPAR  = LSPAR_conv(LSPAR0);
end

DISPLAY_PAR=1;
VERBOSE=0;

if isfield(LSPAR,'C')
  C=LSPAR(1).C;
else
  C=1;
end

starttime = clock;
LW = LW(:);   % jetzt Spaltenvektor
p_a = p_a(:)';
p_b = p_b(:)';
%p_c = p_c(:)';
p_c = 1 - p_a - p_b;

% Remove zero populations
p_a_zero = find(p_a==0);
if (length(p_a_zero) > 0)
  p_a(p_a_zero) = 1e-6;
end
p_b_zero = find(p_b==0);
if (length(p_b_zero) > 0)
  p_b(p_b_zero) = 1e-6;
end
p_c_zero = find(p_c==0);
if (length(p_c_zero) > 0)
  p_c(p_c_zero) = 1e-6;
end
```

```
if any(p_a<0)
  p_a=p_a.^2;
end
if any(p_b<0)
  p_b=p_b.^2;
end
if any(p_c<0)
  p_c=p_c.^2;
end

if tau21<0
  tau21=abs(tau21);
end
if tau32<0
  tau32=abs(tau32);
end
tau12 = tau21 * p_a ./p_b    % Zeilenvektor
tau23 = tau32 * p_b ./p_c    % Zeilenvektor

if isfield(LSPAR(1),'Kex')
  E(1) = LSPAR(1).Kex;
end

if exist('DISPLAY_PAR') & DISPLAY_PAR
  disp('###################################')
  p_a
  p_b
  p_c
  p_a+p_b+p_c
  p_c./p_b
  tau21
  tau32
  disp('###################################')
  %  keyboard
end

TD = points;
T2 = 1 ./ (pi*LW);  % Callaghan p. 195
R2 = diag(1./T2);
DW = 1/(SWH);
AQ = DW * TD;
t = linspace(0,AQ,TD);
const1 = 2*pi*i;
%E = [1,1,1];
%w = w - SWH/TD;        % start at zero, not 1
W = diag(W);

if isfield(LSPAR(1),'Kex')
  if LSPAR(1).Kex~=0 & ~isempty(LSPAR(1).Kex)
    E=[LSPAR(1).Kex,1,1];
  end
elseif isfield(LSPAR(1),'E')
  E=LSPAR(1).E;
  if length(E)~=3
    error('Wrong parameter E, must have length 3.');
  end
else
  E = [1,1,1];
end
```

```
TD_wo_zf = floor(TD/zf);        % TD without additional zf in
                                        % K = p_c/p_b
for m = 1:length(p_a)
  %  K = [1/tau12(m), -1/tau21,     0
  %      -1/tau12(m),  1/tau21+1/tau, -1/tau
  %       0,       -1/tau,      1/tau ];

  %  K = [1/tau12(m), -1/tau21,     0
  %      -1/tau12(m),  1/tau21+1/tau, -1/tau
  %       0,       -1/tau,      1/tau ];

  K = [1/tau12(m), -1/tau21,        0
       -1/tau12(m),  1/tau21+1/tau23(m), -1/tau32
       0,      -1/tau23(m),     1/tau32 ];

  KW = (-K + const1*W - R2);
  P = [p_a(m), p_b(m), p_c(m)]';

  for n = 1 : TD_wo_zf                  % consider zerofilling of orig data
    F_t(n,m) = E * (expm(KW * t(n)) * P);   % FIDs in Spalten
                                            % F_t(TD_wo_zf,m) = 0;
                                            % fill up to TD_wo_zf
  end
end
WDWF = wdwf(:) * ones(1,length(p_a));
F_t_wdwf =  F_t .* WDWF;        % wdwf applied

% keyboard

% 1st point 1/2 removes offset from spc
F_t_wdwf(1,:) = F_t_wdwf(1,:)/2;
F_w = real(fft(F_t_wdwf,TD));       % no fft_shift here
                                        % 2nd arg. TD ==> zerofilling !

SCALEMIN=0;             % not a good idea
if(SCALEMIN)
  MIN = min(min(F_w));
  F_w = F_w - MIN;
end
MAX = max(max(F_w));
F_w = C*F_w/MAX;

% scale with second dimension
scale = diag(scale);
F_w = F_w * scale;

%disp('Total execution time og ftex312 w/o plot):');
%curtime=(clock);
%time = curtime(6) - starttime(6)

%if ndims(LSPAR)==2
%  clear LSPAR
%  LSPAR = LSPAR0;
%end

return
```

```
% ======================================================================
% MKLSPAR - Make LSPAR structure from parameter lineshape input file
% (specified by filename).
% ======================================================================

function LSPAR=mklspar2(filename)
%
%

disp(['Reading parameters from file ' filename '.'])
eval(filename);

% Required fields in setup file
if ~exist('TITLE')
  disp('TITLE missing in lineshape setup file.')
  TITLE=[];
end

if ~exist('USE_DATA')
  disp('USE_DATA missing in lineshape setup file.')
  USE_DATA=0;
end

if USE_DATA==0
  if ~exist('MX1siz')
    disp('MX_1siz missing in lineshape setup file.')
  end
  if ~exist('MX2siz')
    disp('MX2siz missing in lineshape setup file.')
  end
  if ~exist('NO_SPCS')
    disp('NO_SPCS missing in lineshape setup file.')
  end
end

if ~exist('filename')
  disp('filename missing in lineshape setup file.')
end
if ~exist('INVERT')
  disp('INVERT missing in lineshape setup file.')
end
if ~exist('MECH')
  disp('MECH missing in lineshape  setup file.')
end
if ~exist('SWH')
  disp('SWH missing in lineshape setup file.')
end
if ~exist('TD')
  disp('TD missing in lineshape setup file.')
end
if ~exist('Hzppt')
  disp('Hzppt missing in lineshape setup file.')
end
if ~exist('LW')
  disp('LW missing in lineshape setup file.')
end
if ~exist('midmax')
  disp('midmax missing in lineshape setup file.')
end
if ~exist('w')
```

```
  disp('w (omega) missing in lineshape setup file.')
end
if ~exist('zf')
  disp('zf missing in lineshape setup file.')
end
if ~exist('zf0')
  disp('zf0 missing in lineshape setup file.')
end
if ~exist('p_a_0')
  disp('p_a_0 missing in lineshape setup file.')
end
if ~exist('p_b_0')
  disp('p_b_0 missing in lineshape setup file.')
end
if ~exist('p_c_0')
  disp('p_c_0 missing in lineshape setup file.')
end
if ~exist('p_d_0')
  disp('p_d_0 missing in lineshape setup file.')
end
if ~exist('wdwf')
  disp('wdwf missing in lineshape setup file.')
end


if ~exist('DIM')
  disp('DIM missing in lineshape  setup file.')
  DIM=1;
end
if ~exist('filename_2')
  disp('filename_2 missing in lineshape setup file.')
end

if DIM==2
  if ~exist('MECH_2')
    disp('MECH_2 missing in lineshape  setup file.')
  end
  if ~exist('INVERT_2')
    disp('INVERT_2 missing in lineshape  setup file.')
  end
  if ~exist('SWH_2')
    disp('SWH_2 missing in lineshape  setup file.')
  end
  if ~exist('TD_2')
    disp('TD_2 missing in lineshape  setup file.')
  end
  if ~exist('Hzppt_2')
    disp('Hzppt_2 missing in lineshape  setup file.')
  end
  if ~exist('LW_2')
    disp('LW_2 missing in lineshape  setup file.')
  end
  if ~exist('midmax_2')
    disp('midmax_2 missing in lineshape  setup file.')
  end
  if ~exist('w_2')
    disp('w_2 missing in lineshape  setup file.')
  end
  if ~exist('zf0_2')
    disp(' missing in lineshape  setup file.')
  end
```

```
  if ~exist('p_a_0_2')
    disp('p_a_0_2 missing in lineshape  setup file.')
  end
  if ~exist('p_b_0_2')
    disp('p_b_0_2 missing in lineshape  setup file.')
  end
  if ~exist('p_c_0_2')
    disp('p_c_0_2 missing in lineshape  setup file.')
  end
  if ~exist('p_d_0_2')
    disp('p_d_0_2 missing in lineshape  setup file.')
  end
  if ~exist('wdwf_2')
    disp('wdwf_2 missing in lineshape  setup file.')
  end
  if exist('FIT') & FIT==1
    if ~exist('FITDIM')
      FITDIM=[1,1];
    end
  end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% DON't TOUCH ANYTHING AFTER THIS POINT %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

LSPAR(1).TITLE      =  TITLE;

LSPAR(1).USE_DATA   =  USE_DATA;
if USE_DATA==0
  LSPAR(1).MXsiz  = MX1siz;
  LSPAR(2).MXsiz  = MX2siz;

  LSPAR(1).NO_SPCS   = NO_SPCS;
  LSPAR(2).NO_SPCS   = NO_SPCS;
end

LSPAR(1).realdim    = 1;
LSPAR(1).filename   =  filename;;
LSPAR(1).INVERT     =  INVERT;
LSPAR(1).MECH       =  MECH;
LSPAR(1).SWH        =  SWH;
LSPAR(1).TD         =  TD;
LSPAR(1).Hzppt      =  Hzppt;
LSPAR(1).LW         =  LW;
LSPAR(1).midmax     =  midmax;
LSPAR(1).w          =  w;
LSPAR(1).zf         =  zf;
LSPAR(1).zf0        =  zf0;
LSPAR(1).taus_0     =  taus_0;
LSPAR(1).p_a_0      =  p_a_0;
LSPAR(1).p_b_0      =  p_b_0;
LSPAR(1).p_c_0      =  p_c_0;
LSPAR(1).p_d_0      =  p_d_0;
LSPAR(1).wdwf       =  wdwf;
LSPAR(1).DIM        =  DIM;
LSPAR(2).DIM        =  DIM;
LSPAR(1).FIT        =  FIT;
LSPAR(1).FITDIM     =  FITDIM;
%LSPAR.NO_SPCS    =  NO_SPCS;
```

```
% DIM 2
LSPAR(1).realdim    =  2;
LSPAR(2).filename   =  filename_2;
LSPAR(2).INVERT     =  INVERT_2;
LSPAR(2).MECH       =  MECH_2;
LSPAR(2).SWH        =  SWH_2;
LSPAR(2).TD         =  TD_2;
LSPAR(2).Hzppt      =  Hzppt_2;
LSPAR(2).LW         =  LW_2;
LSPAR(2).midmax     =  midmax_2;
LSPAR(2).w          =  w_2;
LSPAR(2).zf0        =  zf0_2;
LSPAR(2).zf         =  zf_2;
if exist('taus_0_2')
  LSPAR(2).taus_0   =  taus_0_2;
else
  LSPAR(2).taus_0   =  taus_0;
end
LSPAR(2).p_a_0      =  p_a_0_2;
LSPAR(2).p_b_0      =  p_b_0_2;
LSPAR(2).p_c_0      =  p_c_0_2;
LSPAR(2).wdwf       =  wdwf_2;

if FIT
  LSPAR(1).OPTIONS    =  OPTIONS;
  LSPAR(2).OPTIONS    =  OPTIONS;
end

% Control parameters
LSPAR(1).PRINT      =  PRINT;
LSPAR(2).PRINT      =  PRINT;
LSPAR(1).PLOT       =  PLOT;
LSPAR(2).PLOT       =  PLOT_2;

% OPTIONAL PARAMETERS
if exist('Kex')
  LSPAR(1).Kex=Kex;
end
if exist('Kex_2')
  LSPAR(2).Kex = Kex_2;
end
if exist('K')
  LSPAR(1).K=K;
end
if exist('K_2')
  LSPAR(2).K=K;
end
if exist('E')
  LSPAR(1).E=E;
end
if exist('E_2')
  LSPAR(2).E=E;
end
if exist('p_L')
  LSPAR(1).p_L=p_L;
end
if exist('P0')
  LSPAR(1).P0=P0;
end
```

```
if exist('DIM2MODE')
  LSPAR(1).DIM2MODE=DIM2MODE;
  LSPAR(2).DIM2MODE=DIM2MODE;
end

return
```

```
% ========================================================================
%                       Test file for NMRKIN
% ========================================================================
% example of lineshape parameters file

global HANDLES

USE_DATA=1;
% IF USE_DATA is 0 you must specify MX1_siz and MX2_siz
% To plot dim 2 data a filename must be specified and DIM must be 2
F=4;
% MX1siz=40*F;
% MX2siz=40*F;
NO_SPCS=13;

TITLE      = 'T91C L21 dim2 using m2121';
filename   = 'MATLAB/nmrkin_cl/Junesim/NMRLAB_DATA_11_L21_d2.mat';  % path and
name of data file
INVERT     = 0;
MECH       = 'm2121_2D';
SWH        = 2000;
TD         = 1024;                  % TD at which wswf was applied
Hzppt      = SWH/TD/F;
REF        = [];                % copy from PROC
LW         = [11 11 11];
midmax     = [6];               % no of row/col of intermediate max
w          = [51 90 115];           % Pos of maxima in points
                        % leave empty for mouse selection
zf         = 1;                 % zeroffilling factor
zf0        = 1;

Kex        = 1.;            % account for external exchange
taus_0     = [1/40 1/20];           % all taus

p_a_0      = [0.99  0.94  0.88  0.77  0.40  0.35  0.24  0.12  0.03  0.008  0.003
0.001  0.0001];

p_b_0      = [0.00  0.03  0.09  0.15  0.30  0.32  0.26  0.25  0.16  0.04   0.010
0.0001  0.0001];
%p_c_0      = [-0.0100 0.1000 0.2000 0.3000 0.3000 0.3000 0.5750 0.71000 0.8680
0.9630 0.9820 0.9990 1.0000];
p_c_0      = 1-p_a_0-p_b_0;

%p_b_0      = [0.001  0.11  0.21  0.31  0.45  0.60  0.70  0.79 0.9 0.999];
%p_c_0      = [0.000  0.00   0.000  0.000  0.000  0.000 0.000 0.000 0.000];
p_d_0      = [0 0 0 0 0 0 0 0 0 0 0 0 0];

wdwf_2     = {'gm',-5,0.2,3255};        % window function type and pars
%gb         = 0.1;
%lb         = -5;

% ========================================================================
%                   Parameters for 2nd dimension
% ========================================================================
DIM        = 2;                 % expect 2D data
TITLE      = 'T91C L21 dim1 using m2121';
filename_2 = 'MATLAB/nmrkin_cl/Junesim/NMRLAB_DATA_11_L21_d1.mat';  %
INVERT_2   = 1;                 % wird im Moment nicht gebraucht
MECH_2     = 'm2121_2D';
```

```
SWH_2      = 6510;
TD_2       = 2048;
Hzppt_2    = SWH_2/TD_2/F;
LW_2       = [21 21 20];
midmax_2   = [6];
w_2        = [82 110 173];              % leave empty for mouse selection
zf0_2      = 1;
zf_2       = 1;
taus_0_2   = taus_0;
p_a_0_2    = p_a_0;
p_b_0_2    = p_b_0;
p_c_0_2    = 1-p_a_0_2-p_b_0_2;
%p_c_0_2    = p_c_0;
%p_d_0_2    = p_d_0;
wdwf       = {'gm',-3,0.15,1600};


% ======================================================================
%                   General control options
% ======================================================================
PRINT      = 1;
PLOT       = 1;
PLOT_2     = 0;

% ======================================================================
%                   fminu / fmins options
% ======================================================================
FIT        = 0;
FITDIM     = [1 0];
OPTIONS=optimset;
OPTIONS=optimset(OPTIONS,'LevenbergMarquardt','off');
%   OPTIONS=optimset(OPTIONS,'LineSearch','cubicpoly');
OPTIONS=optimset(OPTIONS,'LineSearch','quadcubic');
OPTIONS=optimset(OPTIONS,'Diagnostics','on');
%OPTIONS=optimset(OPTIONS,'DiffMaxChange',1e-1);  % Default: 1e-1
%OPTIONS=optimset(OPTIONS,'DiffMinChange',1e-6);  % Default: 1e-8
%OPTIONS=optimset(OPTIONS,'Display','iter');  % 'off,','iter', default:'final'
%OPTIONS=optimset(OPTIONS,'MaxFunEvals',1000);
%OPTIONS=optimset(OPTIONS,'MaxIter',1000);
%OPTIONS=optimset(OPTIONS,'TolCon',1e-3);
%OPTIONS=optimset(OPTIONS,'TolFun',1e-4);
%OPTIONS=optimset(OPTIONS,'TolX',1e-4);


% **********************************************************************
```

```
% ========================================================================
% master function for relaxation dispersion analysis of cL-BABP/S-S data
% ========================================================================

function val = CPMGfull2f_fitfun(par,R2,tau_cp,B0,action)

PLOT=1;
EQN=1;  % 1: approx for many pulses,, 2: approx for slow exch

noFIELDs = length(B0);
B1 = B0(1);
if noFIELDs==2
  B2 = B0(2);
end

% par = real(par);
dw  = abs(real(par(1)));
kex = abs(real(par(2)));
pb  = abs(par(3));
%pb  = 0.02;

R2et1  = par(4);
if noFIELDs==2
  R2et2  = par(5);
end
tau_cp1 = tau_cp(1,:);

data1 = R2(1,:);
if noFIELDs==2
  data2 = R2(2,:);
end
pa = 1 - pb;

if EQN==1   % Millet, 2000, 122, 2867, intermediate-fast
  % for R2a = R2b
  %******** field 1 *************
  PSI1 = (pb * kex - pa * kex)^2 - dw^2 + 4*pa*pb*kex^2;
  CHI1 = 2*dw * (pb * kex - pa * kex);
  % bei Mulder "tau_cp * sqrt(2)", da tau_cp anders definiert ist - nicht
  % fuer Varian Pulsprog
  np1  = tau_cp1*sqrt(2) * ( PSI1 + (PSI1^2 + CHI1^2)^0.5)^0.5;
  nm1  = tau_cp1*sqrt(2) * (-PSI1 + (PSI1^2 + CHI1^2)^0.5)^0.5;
  Dp1  = 0.5 * ( 1 + (PSI1 + 2*dw^2)/((PSI1^2 + CHI1^2)^0.5) );
  Dm1  = 0.5 * (-1 + (PSI1 + 2*dw^2)/((PSI1^2 + CHI1^2)^0.5) );
  y1 = 0.5*( 2*R2et1 + kex - 1./(2*tau_cp1) .* acosh(Dp1 * cosh(np1) - Dm1 *
cos(nm1)));

  if noFIELDs==2
    tau_cp2 = tau_cp(2,:);
    %******** field 2 *************
    dw2  = dw*(B2/B1);
    PSI2 = (pb * kex - pa * kex)^2 - dw2^2 + 4*pa*pb*kex^2;
    CHI2 = 2*dw2 * (pb * kex - pa * kex);
    % bei Mulder "tau_cp * sqrt(2)", da tau_cp anders definiert ist
    % not any more tau_cp/sqrt(2)
    np2  = tau_cp2*sqrt(2) * ( PSI2 + (PSI2^2 + CHI2^2)^0.5)^0.5;
    nm2  = tau_cp2*sqrt(2) * (-PSI2 + (PSI2^2 + CHI2^2)^0.5)^0.5;
    Dp2  = 0.5 * ( 1 + (PSI2 + 2*dw2^2)/((PSI2^2 + CHI2^2)^0.5) );
    Dm2  = 0.5 * (-1 + (PSI2 + 2*dw2^2)/((PSI2^2 + CHI2^2)^0.5) );
```

```
    y2 = 0.5*( 2*R2et2 + kex - 1./(2*tau_cp2) .* acosh(Dp2 * cosh(np2) - Dm2 *
cos(nm2)));
  end
elseif EQN==2    % Tollinger JACS 2001, 123, 11341 for slow exchange
  tau_cp1 = tau_cp(1,:);
  y1 = R2et1 + pa*kex - pa*kex*sin(dw*tau_cp1)./(dw*tau_cp1);
  if noFIELDs==2
    tau_cp2 = tau_cp(2,:);
    dw2  = dw*(B2/B1);
    y2 = R2et1 + pa*kex - pa*kex*sin(dw2*tau_cp2)./(dw2*tau_cp2);
  end
else
  error('EQN must be 1 or 2.')
end

if PLOT
  clf
  plot(1./(4*tau_cp1),data1,'b.')
  hold on
  plot(1./(4*tau_cp1),y1,'b-')
  if noFIELDs==2
    plot(1./(4*tau_cp2),data2,'r.')
    hold on
    plot(1./(4*tau_cp2),y2,'r-')
    hold off
    drawnow
  else
    hold off
    drawnow
  end
end

if action==0
  chi2_1 = sum((data1-y1).^2);
  if noFIELDs==2
    chi2_2 = sum((data2-y2).^2);
    chi2_gesamt = chi2_1 + chi2_2;
  else
    chi2_gesamt = chi2_1;
  end
  val=chi2_gesamt;

elseif action==1
  val(1,:)=y1;
  if noFIELDs==2
    val(2,:)=y2;
  end
end

return
```

Acknowledgments

I would like to acknowledge:

Dr. Laura Ragona, from the NMR Lab., of ISMAC-CNR, Milan where I worked and Prof. Henriette Molinari, from the University of Verona, for their supervision, advice, and contribution they gave me from the earliest stages of this research and for the motivation and encouragement during the whole period of this thesis's work.
Dr. Lucia Zetta for the unique blend of scientific competence and pleasantness.
Dr. Simona Tomaselli for being always available and committed.
The NMR lab. group working with me in Milan: Antonella, Katiuscia, Fulvia, Giulio, Laura Ruth, Maristella, Roberto.
The NMR group of Verona: Michael, Massimo, Serena, Mariapina and Mara.
Prof. U.L Günther for his scientific contribution to this thesis.
The NMR lab. group working at the University of Birmingham: Christian, Kostans, John, Martin, Karen, Alfonso, Sue, Clare and especially Sara.
My parents, Roberto and Rosa, who have been very supportive financially and even more morally, throughout all my studies and without whom, I would not be where I am.
Last but not least, the rest of my family and all my wonderful friends.

Bibliography

1. Mittag, T., et al., *Retinol modulates site-specific mobility of apo-cellular retinol-binding protein to promote ligand binding.* J Am Chem Soc, 2006. **128**(30): p. 9844-8.
2. Gunther, U.L. and B. Schaffhausen, *NMRKIN: simulating line shapes from two-dimensional spectra of proteins upon ligand binding.* J Biomol NMR, 2002. **22**(3): p. 201-9.
3. Strautnieks, S.S., et al., *A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis.* Nat Genet, 1998. **20**(3): p. 233-8.
4. Paulusma, C.C., et al., *A mutation in the human canalicular multispecific organic anion transporter gene causes the Dubin-Johnson syndrome.* Hepatology, 1997. **25**(6): p. 1539-42.
5. Pellicoro, A. and K.N. Faber, *Review article: The function and regulation of proteins involved in bile salt biosynthesis and transport.* Aliment Pharmacol Ther, 2007. **26 Suppl 2**: p. 149-60.
6. Craddock, A.L., et al., *Expression and transport properties of the human ileal and renal sodium-dependent bile acid transporter.* Am J Physiol, 1998. **274**(1 Pt 1): p. G157-69.
7. Tochtrop, G.P., et al., *Energetics by NMR: site-specific binding in a positively cooperative system.* Proc Natl Acad Sci U S A, 2002. **99**(4): p. 1847-52.
8. Dawson, P.A., et al., *The heteromeric organic solute transporter alpha-beta, Ostalpha-Ostbeta, is an ileal basolateral bile acid transporter.* J Biol Chem, 2005. **280**(8): p. 6960-8.
9. Hagenbuch, B. and P.J. Meier, *Molecular cloning, chromosomal localization, and functional characterization of a human liver Na+/bile acid cotransporter.* J Clin Invest, 1994. **93**(3): p. 1326-31.
10. Kullak-Ublick, G.A., et al., *Chlorambucil-taurocholate is transported by bile acid carriers expressed in human hepatocellular carcinomas.* Gastroenterology, 1997. **113**(4): p. 1295-305.
11. Kullak-Ublick, G.A., et al., *Molecular and functional characterization of an organic anion transporting polypeptide cloned from human liver.* Gastroenterology, 1995. **109**(4): p. 1274-82.
12. Alrefai, W.A. and R.K. Gill, *Bile acid transporters: structure, function, regulation and pathophysiological implications.* Pharm Res, 2007. **24**(10): p. 1803-23.
13. Doege, H., et al., *Targeted deletion of FATP5 reveals multiple functions in liver metabolism: alterations in hepatic lipid homeostasis.* Gastroenterology, 2006. **130**(4): p. 1245-58.
14. Maeda, K., et al., *Uptake of ursodeoxycholate and its conjugates by human hepatocytes: role of Na(+)-taurocholate cotransporting polypeptide (NTCP), organic anion transporting polypeptide (OATP) 1B1 (OATP-C), and oatp1B3 (OATP8).* Mol Pharm, 2006. **3**(1): p. 70-7.
15. Ko, J., et al., *Effects of side chain length on ionization behavior and transbilayer transport of unconjugated dihydroxy bile acids: a comparison of nor-chenodeoxycholic acid and chenodeoxycholic acid.* J Lipid Res, 1994. **35**(5): p. 883-92.
16. Rius, M., et al., *Substrate specificity of human ABCC4 (MRP4)-mediated cotransport of bile acids and reduced glutathione.* Am J Physiol Gastrointest Liver Physiol, 2006. **290**(4): p. G640-9.
17. Danielsson, H., ed. *In The Bile Acids: Chemistry, Physiology and Metabolism*. Plenum Press, , ed. P.P.a.K. Nair, D. Vol. 2. 1973: New York. 1–32.
18. Danielsson, H., ed. *The Bile Acids: Chemistry, Physiology and Metabolism* Plenum Press, New York, ed. P.P.a.K. Nair, D. 1971-3: New York
19. Hofmann, A.F., ed. *In Bile Acids and Hepatobiliary Disease* Kluwer, , ed. T. Northfield, Zentler-Munro, P.L. and Jazrawi, R. P. 1999: Boston. 303–332

20. Carey, M.C., ed. *In Phospholipids and Atherosclerosis* Raven Press, ed. P. Avogaro. 1983: New York
21. Hofmann, A.F., *In The Liver: Biology and Pathology* Raven Press, New York, 1994: p. 677.
22. Chiang, J.Y., *Regulation of bile acid synthesis: pathways, nuclear receptors, and mechanisms.* J Hepatol, 2004. **40**(3): p. 539-51.
23. Small, D.M., *In The Bile Acids: Chemistry, Physiology and Metabolism* Plenum Press, New York, 1971. **1**: p. 249–356.
24. *In The Bile Acids: Chemistry, Physiology and Metabolism* Plenum Press,, ed. P.P.a.K. Nair, D. . Vol. 1. 1971, Small, D. M.: New York. 249–356.
25. Small, D.M., ed. Adv. Chem. Ser. Vol. 84. 1968. 31.
26. Carey, M.C. and D.M. Small, *The characteristics of mixed micellar solutions with particular reference to bile.* Am J Med, 1970. **49**: p. 590-608.
27. Ueno, M., ed. *In Structure–Performance Relationships in Surfactants*. ed. K.a.U. Esumi, M. Vol. 147. 1997, Marcel and Dekker,: New York.
28. Wiedmann, T.S. and L. Kamel, *Examination of the solubilization of drugs by bile salt micelles.* J Pharm Sci, 2002. **91**(8): p. 1743-64.
29. Vasile, F., et al., *Solution structure of chicken liver basic fatty acid binding protein.* J Biomol NMR, 2003. **25**(2): p. 157-60.
30. Nichesola, D., et al., *Crystal structure of chicken liver basic fatty acid-binding protein complexed with cholic acid.* Biochemistry, 2004. **43**(44): p. 14072-9.
31. Thompson, J., A. Reese-Wagoner, and L. Banaszak, *Liver fatty acid binding protein: species variation and the accommodation of different ligands.* Biochim Biophys Acta, 1999. **1441**(2-3): p. 117-30.
32. Tomaselli, S., et al., *NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids.* Proteins, 2007. **69**(1): p. 177-91.
33. Haunerland, N.H. and F. Spener, *Fatty acid-binding proteins--insights from genetic manipulations.* Prog Lipid Res, 2004. **43**(4): p. 328-49.
34. Thompson, J., et al., *The liver fatty acid binding protein--comparison of cavity properties of intracellular lipid-binding proteins.* Mol Cell Biochem, 1999. **192**(1-2): p. 9-16.
35. Arighi, C.N., J.P. Rossi, and J.M. Delfino, *Temperature-induced conformational switch in intestinal fatty acid binding protein (IFABP) revealing an alternative mode for ligand binding.* Biochemistry, 2003. **42**(24): p. 7539-51.
36. Eliseo, T., et al., *Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two bile salts revealed by NMR.* Biochemistry, 2007. **46**(44): p. 12557-67.
37. Makishima, M., et al., *Identification of a nuclear receptor for bile acids.* Science, 1999. **284**(5418): p. 1362-5.
38. Cogliati, C., et al., *Disulfide bridge regulates ligand-binding site selectivity in liver bile acid-binding proteins.* Febs J, 2009. **276**(20): p. 6011-23.
39. Creighton, T.E., *Disulphide bonds and protein stability.* Bioessays, 1988. **8**(2): p. 57-63.
40. Dill, K.A., *Theory for the folding and stability of globular proteins.* Biochemistry, 1985. **24**(6): p. 1501-9.
41. Harrison, P.M. and M.J. Sternberg, *The disulphide beta-cross: from cystine geometry and clustering to classification of small disulphide-rich protein folds.* J Mol Biol, 1996. **264**(3): p. 603-23.
42. Thangudu, R.R., et al., *Analycys: a database for conservation and conformation of disulphide bonds in homologous protein domains.* Proteins, 2007. **67**(2): p. 255-61.
43. Richardson, J.S., *The anatomy and taxonomy of protein structure.* Adv Protein Chem, 1981. **34**: p. 167-339.
44. Thornton, J.M., *Disulphide bridges in globular proteins.* J Mol Biol, 1981. **151**(2): p. 261-87.
45. Wouters, M.A., R.A. George, and N.L. Haworth, *"Forbidden" disulfides: their role as redox switches.* Curr Protein Pept Sci, 2007. **8**(5): p. 484-95.

46. Derman, A.I. and J. Beckwith, *Escherichia coli alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm.* J Bacteriol, 1991. **173**(23): p. 7719-22.
47. Kadokura, H., F. Katzen, and J. Beckwith, *Protein disulfide bond formation in prokaryotes.* Annu Rev Biochem, 2003. **72**: p. 111-35.
48. Prinz, W.A., et al., *The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm.* J Biol Chem, 1997. **272**(25): p. 15661-7.
49. Choi, H., et al., *Structural basis of the redox switch in the OxyR transcription factor.* Cell, 2001. **105**(1): p. 103-13.
50. Jakob, U., et al., *Chaperone activity with a redox switch.* Cell, 1999. **96**(3): p. 341-52.
51. Redl, B., et al., *Phage display reveals a novel interaction of human tear lipocalin and thioredoxin which is relevant for ligand binding.* FEBS Lett, 1999. **460**(1): p. 182-6.
52. Liu, J., et al., *Effects of removing a conserved disulfide bond on the biological characteristics of rat lipocalin-type prostaglandin D synthase.* Biochimie, 2008. **90**(11-12): p. 1637-46.
53. Sjoelund, V. and I.A. Kaltashov, *Transporter-to-trap conversion: a disulfide bond formation in cellular retinoic acid binding protein I mutant triggered by retinoic acid binding irreversibly locks the ligand inside the protein.* Biochemistry, 2007. **46**(46): p. 13382-90.
54. Capaldi, S., et al., *A single amino acid mutation in zebrafish (Danio rerio) liver bile acid-binding protein can change the stoichiometry of ligand binding.* J Biol Chem, 2007. **282**(42): p. 31008-18.
55. Ragona, L., et al., *NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein.* J Biol Chem, 2006. **281**(14): p. 9697-709.
56. Glatz, J.F. and J.H. Veerkamp, *A radiochemical procedure for the assay of fatty acid binding by proteins.* Anal Biochem, 1983. **132**(1): p. 89-95.
57. Barlos, K., et al., *Application of 2-chlorotrityl resin in solid phase synthesis of (Leu15)-gastrin I and unsulfated cholecystokinin octapeptide. Selective O-deprotection of tyrosine.* Int J Pept Protein Res, 1991. **38**(6): p. 555-61.
58. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
59. Gunther, U.L., C. Ludwig, and H. Ruterjans, *NMRLAB-Advanced NMR data processing in matlab.* J Magn Reson, 2000. **145**(2): p. 201-8.
60. Aue, W.P., Bartholdi, E. and Ernst, R. R.,, *Two Dimensional Spectroscopy. Application to Nuclear Magnetic Resonance.* J. Chem. Phys., 1976. **64**.
61. Jeener, J. in *Oral presentation in Ampere International Summer School II*. 1971. Basko Polje, Yugoslavia.
62. Braunsschweiler, L.a.E., R. R. , J. Magn. Res., 1983. **53**: p. 521-528.
63. Jeener, J., and Meier, B. M, 1979.
64. Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R., , J. Chem. Phys., 1979. **71**.
65. Cavanagh, J.P., G.; Skelton, J.;, ed. *Protein NMR Spectroscopy*. second edition ed. 2007, Elsevier Academic Press.
66. Morris, G.A.a.F., R. , J. Am. Chem. Soc., 1979. **101**(760).
67. J. Cavanagh, W.J.F., A.G. Palmer, III, N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice*. 1996, USA: Academic Press.
68. Zuiderweg, E.R.P.F., S.W., Biochemistry, 1989. **28**: p. 2387-2391.
69. Marion, D.D., P.C.; Kay, L.E.; Wingfield, P.T.; Bax, A.: Gronenborg, A.M.; Clore, G.M., Biochemistry, 1989. **28**: p. 6150-6156.
70. Marion, D.K., L.E.; Sparks, S.W.; Torchia, D.A.; Bax, A., J. Am. Chem. Soc., 1989. **111**: p. 1515-1517.
71. Oschkinat, H., et al., *Three-dimensional NMR spectroscopy of a protein in solution.* Nature, 1988. **332**(6162): p. 374-6.
72. Bax, A.C., G.M.; Gronenborn, A.M.;, J. Magn. Res., 1990. **88**: p. 425-431.

73. Oleiniczak, E.T.X., R.X.; Fesik, S.W., J. Biomol. NMR, 1992. **2**: p. 655-659.
74. Ikura, M., et al., *An efficient NMR approach for obtaining sequence-specific resonance assignments of larger proteins based on multiple isotopic labeling.* FEBS Lett, 1990. **266**(1-2): p. 155-8.
75. Grzesiek, S.B., A., J. Am. Chem. Soc., 1992. **114**: p. 6291-6293.
76. Sattler. M. et al, 1999.
77. Pawson, T., *Protein modules and signalling networks.* Nature, 1995. **373**(6515): p. 573-80.
78. Korzhnev, D.M., et al., *Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR.* Nature, 2004. **430**(6999): p. 586-90.
79. Hill, R.B.B., C.; DeGrado, W. F.; Palmer, A. G. , *Molecular motions and protein folding: Characterization of the backbone dynamics and folding equilibrium of R2D using 13C NMR spin relaxation.* J. Am. Chem. Soc., 2000. **122**: p. 11610–11619.
80. Eisenmesser, E.Z., et al., *Intrinsic dynamics of an enzyme underlies catalysis.* Nature, 2005. **438**(7064): p. 117-21.
81. Boehr, D.D., et al., *The dynamic energy landscape of dihydrofolate reductase catalysis.* Science, 2006. **313**(5793): p. 1638-42.
82. Popovych, N., et al., *Dynamically driven protein allostery.* Nat Struct Mol Biol, 2006. **13**(9): p. 831-8.
83. Fielding, L., *NMR methods for the determination of protein-ligand dissociation constants.* Curr Top Med Chem, 2003. **3**(1): p. 39-53.
84. Fushman, D., ed. *Determination of Protein Dynamics Using 15N Relaxation Measurements*. BioNMR in Drug Research, ed. O. Zerbe. 283-308.
85. Palmer, A.G., 3rd, C.D. Kroenke, and J.P. Loria, *Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules.* Methods Enzymol, 2001. **339**: p. 204-38.
86. Mulder, F.A., et al., *Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme.* J Am Chem Soc, 2001. **123**(5): p. 967-75.
87. Korzhnev, D.M. and L.E. Kay, *Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding.* Acc Chem Res, 2008. **41**(3): p. 442-51.
88. Oscar Millet, J.P.L., Christopher D. Kroenke, Miquel Pons, and and I. Arthur G. Palmer, *The Static Magnetic Field Dependence of Chemical Exchange. Linebroadening Defines the NMR Chemical Shift Time Scale.* J. Am. Chem. Soc., 2000. **122**: p. 2867-2877.
89. Tollinger, M., et al., *Slow dynamics in folded and unfolded states of an SH3 domain.* J Am Chem Soc, 2001. **123**(46): p. 11341-52.
90. Lunati, E., et al., *Evolution strategy optimization for selective pulses in NMR.* J Magn Reson, 1998. **134**(2): p. 223-35.
91. Shaka, A.J.K., J.; Frenkiel, T.; Freeman, R., J. Magn. Reson., 1983. **52**: p. 335-338.
92. Kay, L.E.K., P.; Saarinen, T, J. Am. Chem. Soc. , 1992. **114**: p. 10663-10665.
93. Schleucher, J.S., M.; Griesinger, C. , ed. *Angew. Chem.*, ed. I.E. Engl. Vol. 32. 1993. 1489-1491
94. Marion, D.I., M.; Tschudin, R.; Bax, A. , J. Magn. Reson., 1989 **85**: p. 393-399.
95. Zerbe O., M.R., Kubinyi H. and Folkers G., BioNMR in DrugResearch, 2002.